

Perbandingan Support Vector Machine dan Model *Transformer* untuk Klasifikasi Sentimen Terkait Redenominasi Rupiah di YouTube

Comparison of Support Vector Machines and Transformer Models for Sentiment Classification Related to the Rupiah Redenomination on YouTube

Andini Putri Amrullah^{*1}, Farid Wajidi², A. Amirul Asnan Cirua³

Program Studi Informatika, Fakultas Teknik, Universitas Sulawesi Barat

E-mail : andiniputriamrullah@gmail.com^{*1}, faridwajidi@unsulbar.ac.id²,

amirulasnancirua@unsulbar.ac.id³

^{*}Corresponding author

Received 17 June 2026; Revised 24 July 2026; Accepted 25 July 2026

Abstrak - Redenominasi rupiah merupakan isu kebijakan moneter yang dapat memengaruhi persepsi masyarakat terhadap nilai mata uang dan perilaku ekonomi. Tanggapan masyarakat terhadap isu ini dapat diamati melalui komentar YouTube. Penelitian terdahulu umumnya membandingkan metode pembelajaran mesin klasik dengan satu model *Transformer*, sedangkan perbandingan SVM, IndoBERT, IndoBERTweet, dan penggabungan probabilitas kedua model *Transformer* masih terbatas. Penelitian ini membandingkan keempat model tersebut dengan menerapkan metode *weighted soft voting* pada IndoBERT dan IndoBERTweet. Data awal berjumlah 1.755 komentar. Setelah melalui tahap pembersihan, *preprocessing*, pelabelan, dan penghapusan duplikat, diperoleh 1.588 komentar. Data dibagi secara terstratifikasi menjadi 70% data pelatihan, 10% data validasi, dan 20% data uji. Teks diproses dalam format *komen_svm* untuk SVM dan *komen_bert* untuk model *Transformer*. Pada *seed* 42, model *hybrid* memperoleh akurasi 76,73% dan Macro F1-score 74,83%, yang merupakan hasil tertinggi pada eksperimen ini. Pada dataset dan konfigurasi penelitian yang digunakan, model *hybrid* memperoleh kinerja tertinggi pada korpus yang memuat variasi bahasa formal dan informal. Namun, hasil penelitian ini belum dapat digeneralisasikan ke seluruh komentar YouTube.

Kata Kunci - Analisis Sentimen, Redenominasi Rupiah, Komentar YouTube, SVM, IndoBERT, IndoBERTweet, *Weighted soft voting*.

Abstract - Rupiah redenomination is a monetary policy issue that may influence public perceptions of currency value and economic behavior. Public responses to this issue can be observed through YouTube comments. Previous studies have generally compared classical machine-learning methods with a single *Transformer* model, while comparisons involving SVM, IndoBERT, IndoBERTweet, and probability-level fusion of both *Transformer* models remain limited. This study compares these four models by applying *weighted soft voting* to IndoBERT and IndoBERTweet. The initial dataset contained 1,755 comments. After cleaning, preprocessing, labeling, and duplicate removal, 1,588 comments remained. The dataset was divided using stratified splitting into 70% training, 10% validation, and 20% test data. The text was processed as *komen_svm* for SVM and *komen_bert* for *Transformer*-based models. At *seed* 42, the hybrid model achieved 76.73% accuracy and a 74.83% Macro F1-score, representing the highest performance in this experiment. Within the dataset and experimental configuration used in this study, the hybrid model achieved the highest performance on a corpus containing formal and informal language variations. However, these findings cannot be generalized to all YouTube comments.

Keywords - Sentiment Analysis, Rupiah Redenomination, YouTube Comments, SVM, IndoBERT, IndoBERTweet, *Weighted soft voting*.

1. PENDAHULUAN

Pembahasan mengenai redenominasi rupiah kembali menjadi isu kebijakan moneter yang sangat relevan dan layak untuk dikaji. Inti dari kebijakan ini adalah penyederhanaan nilai pecahan mata uang tanpa mengubah nilai riil atau daya beli masyarakat [1]. Sebagai gambaran, melalui proses redenominasi, nilai nominal Rp10.000 akan disederhanakan dengan menghilangkan tiga angka nol, sehingga menjadi Rp10, tanpa mengurangi nilai atau daya beli dari sepuluh ribu rupiah tersebut [2]. Meskipun secara teknis tampak sederhana, redenominasi memiliki implikasi psikologis yang signifikan bagi masyarakat dalam hal pemahaman terhadap harga dan nilai uang [3]. Di kalangan pelaku usaha, kebijakan ini memang diyakini mampu menyederhanakan pembukuan keuangan. Namun, kekhawatiran terkait munculnya biaya adaptasi, risiko pembulatan harga, serta perubahan persepsi konsumen terhadap daya beli tetap menjadi hal yang perlu diperhatikan [4]. Hal ini menyoroti pentingnya menganalisis opini warganet untuk memetakan respons masyarakat terhadap kebijakan ini dengan lebih baik.

Di era digital saat ini, media sosial telah menjadi platform utama bagi masyarakat untuk menyuarakan pendapat mereka [5]. Media sosial memuat berbagai aktivitas pengguna yang menghasilkan data dalam jumlah besar (*big data*), sehingga dapat dimanfaatkan untuk menganalisis perilaku dan opini masyarakat [6]. Penggunaan *Natural Language Processing (NLP)* terus berkembang untuk mengungkap pola dalam opini, emosi, dan respons pengguna terhadap isu-isu digital. Salah satu platform digital yang memiliki pengaruh besar dalam penyebaran informasi dan pembentukan opini publik di Indonesia adalah YouTube [7]. Karakteristik platform video ini menjadikannya media yang ideal untuk mempelajari reaksi masyarakat terhadap isu-isu terkini. Melalui kolom komentar, pengguna dapat secara terbuka mengungkapkan berbagai tanggapan, termasuk kritik, dukungan, dan pertanyaan. Namun, karakteristik bahasa dalam komentar YouTube bersifat informal dengan penggunaan bahasa gaul, emoji, serta tanda baca yang tidak baku, yang dapat mempersulit proses analisis sentimen dan pemahaman makna secara tepat [8].

Dalam hal klasifikasi sentimen, Support Vector Machine (SVM) tetap menjadi pilihan populer karena ketahanannya terhadap data teks berdimensi tinggi, terutama bila dikombinasikan dengan fitur TF-IDF [9]. Secara teknis, algoritma ini bekerja dengan menentukan *hyperplane* yang optimal untuk memisahkan berbagai kelas di ruang fitur [10]. Namun demikian, sebagai model klasik, SVM sering kali kesulitan menangkap konteks kalimat yang lebih dalam seperti ironi atau makna implisit [11]. Di sinilah model berbasis *Transformer* menawarkan keunggulan yang jelas [12], [13]. IndoBERT telah banyak digunakan untuk memahami konteks bahasa Indonesia dalam berbagai studi sentimen [14], sedangkan IndoBERTweet dianggap lebih optimal untuk menangani teks media sosial, yang cenderung menggunakan bahasa informal [15].

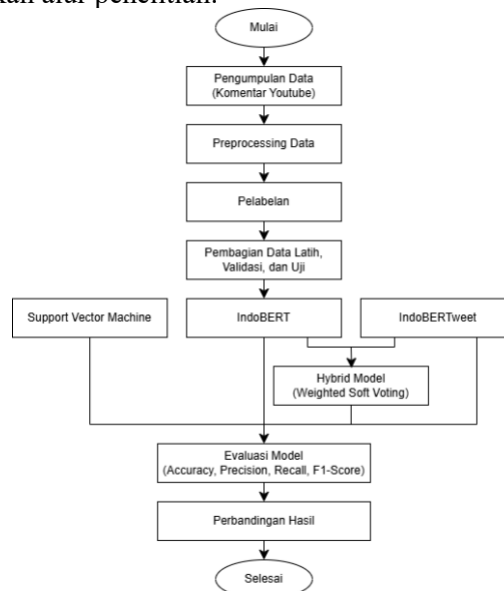
Penelitian sebelumnya telah secara luas menerapkan SVM, IndoBERT, dan IndoBERTweet, menunjukkan kinerja yang baik dalam analisis sentimen di berbagai bidang. Beberapa model ini telah diterapkan tidak hanya pada isu politik [16], tetapi juga pada ulasan layanan kesehatan [17], layanan akademik mahasiswa [18], serta tanggapan publik terhadap kebijakan kendaraan listrik yang dianalisis melalui komentar di YouTube [19]. Pada penelitian [19], pendekatan berbasis TF-IDF dan IndoBERTweet dibandingkan untuk menganalisis sentimen komentar YouTube mengenai kebijakan insentif kendaraan listrik, sedangkan penelitian [20] membandingkan RoBERTa dan Multinomial Naive Bayes pada komentar video pendidikan YouTube dari berbagai saluran. Kedua penelitian tersebut menunjukkan bahwa model berbasis *Transformer* mampu memberikan kinerja yang kuat pada karakteristik data yang digunakan. Namun, studi [19] hanya mengevaluasi satu model *Transformer* IndoBERTweet yang dibandingkan dengan pendekatan berbasis TF-IDF, sedangkan studi [20] membandingkan RoBERTa dan Multinomial Naive Bayes tanpa melibatkan model yang dikembangkan khusus untuk bahasa Indonesia atau pendekatan *hybrid*. Selain itu, kedua studi tersebut menggunakan dataset, langkah *pre-processing*, distribusi kelas, dan protokol evaluasi yang berbeda, sehingga hasil yang dilaporkan tidak dapat dibandingkan secara langsung. Oleh karena itu, masih terdapat keterbatasan dalam

memahami bagaimana kinerja model klasik, model *Transformer*, dan model *hybrid* ketika dievaluasi pada dataset yang sama dan di bawah kondisi uji coba yang sama.

Mengingat kesenjangan tersebut, penelitian ini membandingkan kinerja SVM, IndoBERT, IndoBERTweet, dan model *hybrid* IndoBERT–IndoBERTtweet dalam mengklasifikasi sentimen komentar dari satu video YouTube mengenai redenominasi rupiah. Korpus penelitian terdiri atas 1.588 komentar yang diberi label positif, negatif, dan netral. Pendekatan *weighted soft voting* digunakan untuk menggabungkan probabilitas prediksi IndoBERT dan IndoBERTtweet. Metode tersebut sebelumnya telah diterapkan dengan menggabungkan Multinomial Naïve Bayes dan SVM dalam analisis sentimen RUU TNI di platform X. Model ensemble tersebut mencapai akurasi 96,00%, lebih tinggi daripada SVM sebesar 95,34% dan Naïve Bayes sebesar 93,67% [21]. Namun, penelitian tersebut menggunakan model berbasis TF-IDF pada data dari platform X, sedangkan penelitian ini menggabungkan dua model Transformer berbahasa Indonesia dan membandingkannya dengan SVM pada komentar YouTube. Kontribusi penelitian ini meliputi penyusunan korpus berlabel yang terdiri dari 1.588 komentar, perbandingan empat konfigurasi model menggunakan label, pembagian data, data uji, dan metrik evaluasi yang sama, serta evaluasi model *hybrid* IndoBERT–IndoBERTtweet berdasarkan *weighted soft voting* dengan tahap *preprocessing* yang disesuaikan dengan masing-masing jenis model. Pada dataset dan konfigurasi yang digunakan, model *hybrid* mencapai kinerja yang lebih tinggi daripada model tunggal yang diuji. Namun, penelitian ini masih terbatas pada satu video, distribusi kelas yang belum sepenuhnya seimbang, konfigurasi tertentu, dan tidak mencakup pengujian lintas waktu, video, kanal, dan topik sehingga hasilnya belum dapat digeneralisasikan secara luas.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif eksperimental untuk membandingkan kinerja beberapa model klasifikasi sentimen pada komentar YouTube yang berkaitan dengan isu redenominasi rupiah. Model yang digunakan meliputi Support Vector Machine, IndoBERT, IndoBERTtweet, serta model *hybrid* IndoBERT–IndoBERTtweet yang didasarkan pada metode *weighted soft voting*. Tahapan penelitian meliputi pengumpulan data, pelabelan sentimen, *preprocessing* teks, pembagian data, pelatihan model, dan evaluasi kinerja sebagaimana ditunjukkan pada Gambar 1 yang merupakan alur penelitian.



Gambar 1. Alur Penelitian

2.1. Pengumpulan Data

Data penelitian diperoleh dari komentar pada salah satu video YouTube yang ditemukan melalui pencarian menggunakan kata kunci “redenominasi rupiah”. Video yang dipilih secara purposive adalah video berjudul “Purbaya Targetkan RUU Redenominasi Rupiah Rampung Tahun 2027” dengan ID video ‘c-zPRBgJg84’, karena secara langsung membahas isu redenominasi rupiah dan memiliki jumlah komentar publik yang cukup untuk analisis sentimen. Pengumpulan data dilakukan menggunakan YouTube Data API v3 pada 26 Februari 2026, dengan mengambil semua komentar yang tersedia sejak video tersebut dipublikasikan hingga saat pengumpulan data, termasuk komentar utama dan balasan yang dapat diakses publik. Data yang diterima berupa komentar berbahasa Indonesia, termasuk penggunaan bahasa informal, singkatan, bahasa gaul, variasi ejaan, emoji, dan campuran istilah asing yang tetap dapat dipahami dalam konteks bahasa Indonesia. Data tanpa teks, duplikat, komentar yang tidak relevan dengan isu redenominasi rupiah, atau yang tidak memiliki konten linguistik yang cukup untuk analisis sentimen dikecualikan dari kumpulan data. Proses pengumpulan data menghasilkan 1.755 komentar mentah, setelah validasi dan *preprocessing*, 1.588 komentar digunakan dalam eksperimen. Penggunaan kata kunci pencarian, video ID, dan pencatatan waktu pengambilan data memungkinkan prosedur pengumpulan data direplikasi, meskipun jumlah komentar dapat berubah apabila data diambil pada waktu yang berbeda.

2.2. Preprocessing Data

Tahap *preprocessing* dilakukan untuk membersihkan dan mempersiapkan teks komentar sebelum digunakan dalam proses klasifikasi. *Preprocessing* umum meliputi normalisasi huruf besar-kecil, penghapusan URL, sebutan, tagar, angka, dan tanda baca, serta normalisasi spasi kosong. Selain itu, kata-kata non-standar, singkatan, bahasa gaul, kesalahan ketik, dan variasi ejaan dinormalisasi menggunakan kamus normalisasi dan aturan tambahan.

Hasil *preprocessing* kemudian dibagi menjadi dua aliran. Aliran pertama adalah *komen_svm*, yang berisi teks yang digunakan untuk model Support Vector Machine. Dalam alur ini, teks diproses lebih lanjut dengan menghapus emoji atau emotikon, menghapus *stopword*, dan menerapkan *stemming* menggunakan Sastrawi. Alur kedua adalah *komen_bert*, yang terdiri dari teks yang digunakan untuk IndoBERT dan IndoBERTweet. Dalam alur ini, teks tidak melalui proses penghapusan *stopword* atau *stemming* untuk mempertahankan struktur kalimat, dan emoji atau emotikon dipertahankan untuk menjaga informasi ekspresif dalam komentar.

2.3. Pelabelan Data

Pelabelan sentimen terhadap objek redenominasi dilakukan ke dalam tiga kategori, yaitu positif, negatif, dan netral. Label positif diberikan pada komentar yang menunjukkan dukungan atau persetujuan terhadap redenominasi, label negatif pada komentar yang menunjukkan penolakan, kritik, atau kekhawatiran, sedangkan label netral pada komentar yang bersifat informatif atau tidak menunjukkan sikap evaluatif yang jelas. Tiga anotator diberikan pedoman operasional dan contoh pelabelan, kemudian melakukan anotasi secara *blind* dan independen tanpa mengetahui hasil anotator lainnya. Hasil anotasi dibandingkan untuk menghitung jumlah kesepakatan penuh (3:0) dan kesepakatan mayoritas (2:1), serta mengukur reliabilitas antar-anotator menggunakan Fleiss’ kappa. Label akhir ditentukan berdasarkan *majority voting*. Komentar yang mengandung *mixed sentiment* diberi label berdasarkan sikap yang paling dominan terhadap redenominasi, sedangkan kasus ambigu ditinjau kembali berdasarkan pedoman dan hasil anotasi untuk menentukan label akhir. Selanjutnya, sisa data dilabeli oleh peneliti berdasarkan pedoman yang telah diberikan. Contoh hasil pelabelan data ditunjukkan pada Tabel 1 berikut.

Tabel 1. Contoh Pelabelan Data

Label	Contoh Komentar
Positif	“Saya setuju terhadap kebijakan ini”
Negatif	“Tidak perlu dilakukan kebijakan ini”
Netral	“Redenominasi itu mengurangi nol, bukan nilai uang”

2.4. Data Splitting

Sebelum pemisahan data, entri duplikat identik dihapus berdasarkan kombinasi kolom *komen_svm*, *komen_bert*, dan label untuk mencegah data yang sama tersebar pada subset berbeda. Komentar utama dan komentar balasan tetap diperlakukan sebagai pengamatan terpisah karena tidak dilakukan pengelompokan berdasarkan hubungan komentar maupun kemiripan semantik. Dataset kemudian dibagi menggunakan fungsi *train_test_split* pada pustaka Scikit-learn dengan *random_state*=42 dan parameter *stratify* berdasarkan label sentimen. Pemisahan dilakukan dalam dua tahap sehingga diperoleh 70% data pelatihan, 10% data validasi, dan 20% data uji. Data pelatihan digunakan untuk melatih model, sedangkan data validasi digunakan untuk memilih konfigurasi dan bobot yang optimal. Secara khusus untuk SVM, setelah konfigurasi terbaik diperoleh, model dilatih ulang menggunakan kombinasi data pelatihan dan data validasi. Data uji digunakan hanya untuk evaluasi akhir model.

2.5. Pemodelan Klasifikasi

Eksperimen klasifikasi menggunakan empat model: Support Vector Machine (SVM), IndoBERT, IndoBERTweet, dan model *hybrid* IndoBERT–IndoBERTweet. Konfigurasi masing-masing model serta bobot model *hybrid* dipilih berdasarkan nilai Macro F1-score pada data validasi. Setelah konfigurasi terbaik diperoleh, model SVM dilatih ulang menggunakan kombinasi data pelatihan dan validasi, kemudian dievaluasi pada data uji. Untuk model SVM, fitur diekstraksi dari kolom *komen_svm* dan dikonversi menjadi representasi numerik menggunakan TF-IDF (Term Frequency–Inverse Document Frequency). Bobot TF-IDF dirumuskan dalam Persamaan (1).

$$W_{t,d} = tf_{t,d} \times \log \left(\frac{N}{df_{t,d}} \right) \quad (1)$$

Keterangan:

- $W_{t,d}$: Bobot kata t pada dokumen d .
- $tf_{t,d}$: Frekuensi kemunculan kata t pada dokumen d .
- N : Jumlah seluruh dokumen dalam dataset.
- $df_{t,d}$: Jumlah dokumen yang mengandung kata t .

Model SVM menggunakan gabungan fitur *word* TF-IDF dan *character* TF-IDF melalui FeatureUnion dengan bobot masing-masing 1,0. *Word* TF-IDF menggunakan *n-gram* 1–2, *min_df*=3, *max_df*=1,0, *max_features*=None, *sublinear_tf*=True, dan normalisasi L2. Sementara itu, *character* TF-IDF menggunakan analyzer *char_wb*, *n-gram* 3–5, *min_df*=2, *max_df*=0,9, *max_features*=12.000, *sublinear_tf*=False, dan normalisasi L2. Kedua fitur menghasilkan 10.458 fitur gabungan. Seleksi fitur dilakukan menggunakan SelectKBest dengan fungsi chi-square dan nilai $k = 12.000$. Karena jumlah fitur aktual lebih kecil dari nilai k , seluruh fitur digunakan. Model LinearSVC menerapkan strategi *one-vs-rest*, *class_weight*=*balanced*, maksimum 30.000 iterasi, dan nilai $C = 0,3$ yang dipilih berdasarkan Macro F1 pada data validasi. Fungsi keputusan SVM dirumuskan pada Persamaan (2).

$$f(x) = w^T x + b \quad (2)$$

Keterangan:

- $f(x)$: Fungsi Keputusan model terhadap data masukan x .
 w : Vektor bobot yang menentukan arah *hyperplane*.
 x : Vektor fitur hasil pembobotan TF-IDF.
 b : Bias.

Dalam klasifikasi multi-kelas, kelas yang diprediksi ditentukan berdasarkan nilai fungsi keputusan terbesar untuk setiap kelas, seperti yang ditunjukkan pada persamaan berikut:

$$\hat{y} = \arg \max_k (w_k^T x + b_k) \quad (3)$$

Keterangan:

- $\hat{y}$: Kelas sentimen yang diprediksi.
 k : Indeks kelas sentimen, yaitu positif, negatif, atau netral.
 w_k : Vektor bobot untuk kelas ke- k .
 b_k : Bias untuk kelas ke- k .

Model IndoBERT dan IndoBERTtweet menggunakan data dari kolom *komen_bert*, IndoBERT menggunakan *checkpoint* indobenchmark/indobert-base-p1, sedangkan IndoBERTtweet menggunakan indolem/indoberttweet-base-uncased. Eksperimen dijalankan menggunakan Python 3.12.13, Transformers 5.13.1, dan PyTorch 2.11.0+cu128 pada GPU Tesla T4. Tokenisasi dilakukan menggunakan *tokenizer* masing-masing model dengan *truncation* dan *padding* hingga panjang maksimum 128 token. Proses *fine-tuning* menggunakan *optimizer* AdamW (adamw_torch) dengan beta-1 sebesar 0,9, beta-2 sebesar 0,999, epsilon sebesar 1×10^{-8} , *weight decay* sebesar 0,01, *linear learning-rate scheduler*, dan *warm-up ratio* sebesar 0,10. Ukuran *batch* pelatihan ditetapkan sebesar 16 dan *batch* evaluasi sebesar 32. *Hidden dropout*, *attention dropout*, dan *classifier dropout* efektif masing-masing sebesar 0,1. Evaluasi dan penyimpanan *checkpoint* dilakukan pada setiap *epoch* dengan *early stopping patience* sebesar dua *epoch*. *Checkpoint* terbaik dipilih berdasarkan nilai Macro F1-score pada data validasi. Apabila terdapat nilai Macro F1-score yang sama, akurasi digunakan sebagai kriteria pembandingan.

Hyperparameter tuning dilakukan secara terpisah pada IndoBERT dan IndoBERTtweet menggunakan *seed* 42. Parameter yang diuji meliputi *learning rate* 1×10^{-5} , 2×10^{-5} , dan 3×10^{-5} , serta maksimum *epoch* 3, 4, dan 5. Data uji tidak digunakan dalam pemilihan konfigurasi. Konfigurasi terbaik IndoBERT diperoleh pada *learning rate* 3×10^{-5} dan maksimum 5 *epoch* dengan waktu pelatihan 140,38 detik. Sementara itu, IndoBERTtweet memperoleh konfigurasi terbaik pada *learning rate* 3×10^{-5} dan maksimum 4 *epoch* dengan waktu pelatihan 216,18 detik. Konfigurasi terpilih selanjutnya diuji menggunakan *seed* 42, 52, dan 62 untuk mengetahui kestabilan hasil.

IndoBERT digunakan untuk merepresentasikan konteks bahasa Indonesia secara umum, sedangkan IndoBERTtweet digunakan karena model tersebut dikembangkan dari korpus media sosial. Pada model *Transformer*, probabilitas kelas diperoleh melalui fungsi *softmax*, seperti yang ditunjukkan pada Persamaan (4).

$$P(y|x) = \text{softmax}(Wh + b) \quad (4)$$

Keterangan:

- $P(y|x)$: Probabilitas kelas sentimen y terhadap teks masukan x .
 h : Vektor representasi teks
 W : Matriks bobot pada lapisan klasifikasi.
 b : Bias pada lapisan klasifikasi.
 softmax : Mengubah keluaran model menjadi probabilitas pada setiap kelas.

Model *hybrid* IndoBERT–IndoBERTweet dikembangkan menggunakan metode *weighted soft voting*, yaitu dengan menggabungkan probabilitas prediksi yang dihasilkan oleh kedua model. Dalam metode ini, setiap model diberi bobot yang mencerminkan kontribusinya terhadap prediksi akhir. Probabilitas akhir model *hybrid* dihitung menggunakan Persamaan (5).

$$P_{\text{hybrid}} = w \times P_{\text{IndoBERT}} + (1 - w) \times P_{\text{IndoBERTtweet}} \quad (5)$$

Keterangan:

- P_{hybrid} : Probabilitas akhir dari model *hybrid*.
 P_{IndoBERT} : Probabilitas prediksi dari model IndoBERT.
 $P_{\text{IndoBERTtweet}}$: Probabilitas prediksi dari model IndoBERTtweet.
 w : Bobot untuk model IndoBERT.
 $(1 - w)$: Bobot untuk model IndoBERTtweet.

Penentuan bobot model *hybrid* dilakukan menggunakan data validasi. Bobot IndoBERT (w) diuji pada rentang 0,10–0,90 dengan interval 0,05, sedangkan bobot IndoBERTtweet ditetapkan sebesar $(1 - w)$. Pada setiap *seed*, probabilitas kedua model digabungkan menggunakan Persamaan (5), kemudian dievaluasi berdasarkan Macro F1 pada data validasi. Bobot dengan nilai Macro F1 tertinggi dipilih dan selanjutnya digunakan untuk menghasilkan prediksi pada data uji. Data uji tidak dilibatkan dalam proses pemilihan bobot.

2.6. Evaluasi Model

Evaluasi model dilakukan untuk mengukur kemampuan masing-masing model dalam mengklasifikasikan komentar YouTube ke dalam kategori positif, negatif, dan netral. Evaluasi ini dilakukan menggunakan data uji yang tidak digunakan baik dalam proses pelatihan maupun validasi. Metrik yang digunakan dalam penelitian ini meliputi *Accuracy*, *Precision*, *Recall*, *F1-score*, dan *Confusion matrix*. Pada penelitian ini, nilai *Precision*, *Recall*, dan *F1-score* keseluruhan dihitung menggunakan metode rata-rata makro (*macro average*), yaitu dengan menghitung rata-rata nilai metrik yang diperoleh dari masing-masing kelas sentimen.

1. *Accuracy*: Metrik yang menunjukkan rasio prediksi yang benar terhadap jumlah total data yang diuji.
2. *Precision*: Metrik yang menunjukkan ketepatan model dalam memprediksi kelas tertentu, didefinisikan sebagai rasio prediksi yang benar untuk kelas tersebut terhadap jumlah total data yang diprediksi sebagai kelas tersebut.
3. *Recall*: Metrik yang menunjukkan kemampuan model untuk mengidentifikasi data dalam suatu kelas, khususnya rasio data yang diprediksi dengan benar terhadap semua data aktual dalam kelas tersebut.
4. *F1-score*: Metrik yang mencerminkan keseimbangan antara presisi dan *recall* dengan menggunakan rata-rata harmonik dari kedua metrik tersebut.
5. *Confusion matrix*: Matriks yang digunakan untuk menampilkan perbandingan antara label aktual dan label yang diprediksi model, sehingga mengungkapkan jumlah data yang diklasifikasikan dengan benar dan salah di setiap kelas.

3. HASIL DAN PEMBAHASAN

3.1. Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan komentar pengguna YouTube yang diperoleh melalui proses *scraping* pada hari Kamis, 26 Februari 2026 dari video YouTube berjudul “Purbaya Targetkan RUU Redenominasi Rupiah Rampung Tahun 2027” yang diunggah oleh saluran Tribunnews pada November 2025. Data yang dianalisis mencakup seluruh interaksi digital penonton yang terakumulasi sejak video dipublikasikan hingga waktu pengambilan data menggunakan YouTube Data API v3. Dari proses tersebut, diperoleh data mentah sebanyak 1.755

komentar yang disimpan dalam format Comma-Separated Values (CSV) dengan mengekstrak empat atribut utama, yaitu *author*, *comment*, *likes*, dan *is_reply*, di mana karakteristik data hasil pengumpulan tersebut secara ringkas disajikan pada Tabel 2.

Tabel 2. Hasil Pengumpulan Data

<i>author</i>	<i>comment</i>	<i>likes</i>	<i>Is_reply</i>
@11122	Benar pak memang rakyat mengharap redenominasi itu sangat setuju	0	FALSE

3.2. Preprocessing Data

Tahap *preprocessing* menghasilkan dua format teks *komen_svm* dan *komen_bert*. Kolom *komen_svm* berisi teks yang lebih ringkas karena telah melalui proses normalisasi, penghapusan *stopword*, dan *stemming*. Sementara itu, kolom *komen_bert* mempertahankan struktur kalimat dan emoji untuk menjaga konteks komentar guna proses tokenisasi pada model IndoBERT dan IndoBERTweet. Kumpulan data, yang awalnya terdiri dari 1.755 data, berkurang setelah tahap *preprocessing* karena adanya data kosong, duplikat, atau data yang tidak memenuhi kriteria analisis. Setelah tahap *preprocessing* selesai, jumlah data yang digunakan dalam penelitian ini menjadi 1.588.

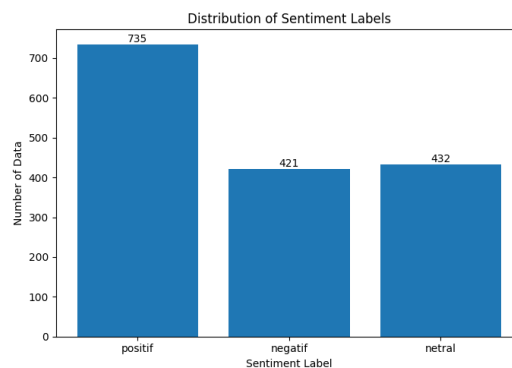
Tabel 3. Hasil Preprocessing

Komentar Asli	Hasil <i>Preprocessing</i>	
	<i>komen_svm</i>	<i>komen_bert</i>
Setuju bgt dgn redenominasi, tapi semoga ga terjadi inflasi krn faktor psikologis pasar 👉👉👉	setuju banget redenominasi semoga tidak jadi inflasi faktor psikologis pasar	setuju banget dengan redenominasi tapi semoga tidak terjadi inflasi karena faktor psikologis pasar 👉👉👉

Berdasarkan Tabel 3, hasil untuk *komen_svm* lebih sederhana karena disesuaikan dengan persyaratan model SVM, yang menggunakan fitur TF-IDF. Sementara itu, hasil untuk *komen_bert* mempertahankan struktur kalimat sehingga model berbasis *Transformer* dapat memahami konteks teks dengan lebih baik.

3.3. Pelabelan Data

Berdasarkan hasil pelabelan, sebanyak 1.588 komentar dikategorikan ke dalam tiga kelas sentimen, yaitu 735 komentar positif, 421 komentar negatif, dan 432 komentar netral. Sebagaimana ditunjukkan pada Gambar 2, sentimen positif merupakan kelas dengan jumlah terbanyak. Temuan ini menunjukkan bahwa komentar pada video YouTube yang digunakan sebagai sumber dataset dalam penelitian ini cenderung didominasi oleh tanggapan positif terhadap isu redenominasi rupiah.



Gambar 2. Distribusi Label

3.4. Data Splitting

Setelah proses validasi ulang dan penghapusan duplikat berdasarkan kolom *komen_svm*, *komen_bert*, dan label, jumlah data yang digunakan dalam eksperimen model adalah 1.588 komentar. Data dibagi menggunakan teknik stratified splitting dengan rasio 70% data pelatihan, 10% data validasi, dan 20% data pengujian. Teknik ini mempertahankan proporsi setiap kelas sentimen dalam setiap subset sehingga distribusi kelas di seluruh dataset pelatihan, validasi, dan pengujian tetap relatif konsisten. Sebagaimana ditunjukkan pada Tabel 4.

Tabel 4. Pembagian Data

Dataset	Negatif	Netral	Positif	Total
Pelatihan	295	302	514	1111
Validasi	42	43	74	159
Pengujian	84	87	147	318
Total	421	432	735	1588

Subset data uji yang sama digunakan pada seluruh model sehingga hasil evaluasi akhir dapat dibandingkan menggunakan data yang sama. Data pelatihan digunakan untuk melatih model, data validasi dimanfaatkan untuk menentukan konfigurasi atau model terbaik selama proses pelatihan, sedangkan data pengujian digunakan hanya pada tahap evaluasi akhir untuk mengukur kemampuan generalisasi model terhadap data yang belum pernah dilihat sebelumnya.

3.5. Hasil Evaluasi Model

Evaluasi performa dilakukan terhadap empat model, yaitu Support Vector Machine, IndoBERT, IndoBERTweet, dan model *hybrid* IndoBERT–IndoBERTweet. Pengujian diukur menggunakan empat metrik utama, yaitu accuracy, Macro Precision, Macro Recall, dan Macro F1-score. Pada model *hybrid*, bobot IndoBERT dan IndoBERTweet dipilih berdasarkan nilai Macro F1 tertinggi pada data validasi untuk setiap *seed*. Bobot optimal yang diperoleh ditunjukkan pada Tabel 5.

Tabel 5. Bobot Optimal Model Hybrid

Seed	Bobot IndoBERT	Bobot IndoBERTweet
42	0,35	0,65
52	0,20	0,80
62	0,25	0,75

Berdasarkan evaluasi data validasi, bobot optimal untuk model *hybrid* berbeda-beda untuk setiap *seed*. Untuk *seed* 42, bobot untuk IndoBERT dan IndoBERTweet masing-masing adalah 0,35 dan 0,65. Untuk *seed* 52, bobot yang dipilih adalah 0,20 dan 0,80, sedangkan untuk *seed* 62, bobotnya adalah 0,25 dan 0,75. Bobot-bobot ini kemudian digunakan untuk menghasilkan prediksi pada data uji tanpa melakukan penyesuaian apa pun menggunakan data uji tersebut.

Tabel 6. Hasil Evaluasi Model pada seed 42

Model	Accuracy	Macro Precision	Macro Recall	Macro F1-Score
Support Vector Machine	0,6792	0,6476	0,6483	0,6479
IndoBERT	0,7327	0,7204	0,7135	0,7130
IndoBERTweet	0,7390	0,7274	0,7015	0,7093
Hybrid IndoBERT–IndoBERTweet	0,7673	0,7606	0,7429	0,7483

Pada *seed* 42, model *hybrid* memperoleh kinerja tertinggi dengan akurasi 0,7673 dan Macro F1-score 0,7483. Untuk menguji kestabilan hasil, IndoBERT, IndoBERTweet, dan model *hybrid* dijalankan menggunakan *seed* 42, 52, dan 62. Ringkasan hasilnya ditunjukkan pada Tabel 7.

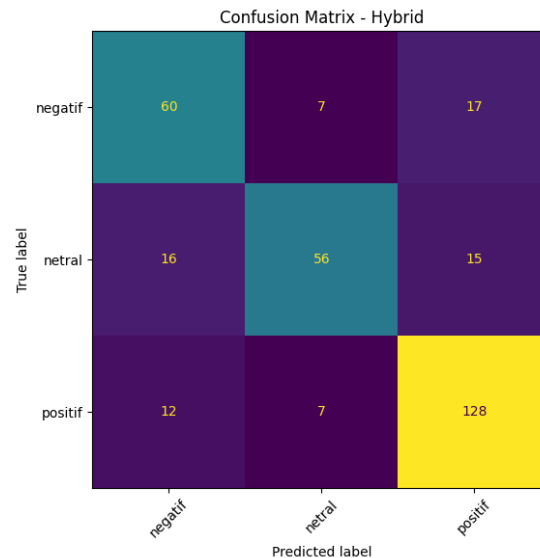
Tabel 7. Hasil Evaluasi Multi-seed

Model	Accuracy Rerata $\pm$ SD	Macro F1 Rerata $\pm$ SD
IndoBERT	0,7338 $\pm$ 0,0018	0,7085 $\pm$ 0,0074
IndoBERTweet	0,7453 $\pm$ 0,0063	0,7192 $\pm$ 0,0097
Hybrid IndoBERT–IndoBERTweet	0,7600 $\pm$ 0,0065	0,7374 $\pm$ 0,0096

Berdasarkan tiga *seed* yang digunakan dalam penelitian ini, model *hybrid* memperoleh nilai rata-rata akurasi dan Macro F1-score tertinggi dibandingkan IndoBERT dan IndoBERTweet. Hasil tersebut menunjukkan bahwa pada dataset dan konfigurasi penelitian ini, penggabungan probabilitas kedua model cenderung memberikan kinerja yang lebih baik dan relatif konsisten. Namun, temuan ini masih terbatas pada dataset, konfigurasi, dan jumlah *seed* yang digunakan dalam penelitian ini.

3.6. Analisis Confusion matrix

Analisis *confusion matrix* dilakukan terhadap model *hybrid*. Model tersebut mengklasifikasi dengan benar 60 dari 84 komentar negatif, 56 dari 87 komentar netral, dan 128 dari 147 komentar positif. Pada kelas negatif, 7 komentar diprediksi sebagai netral dan 17 sebagai positif. Pada kelas netral, 16 komentar diprediksi sebagai negatif dan 15 sebagai positif. Sementara itu, pada kelas positif, 12 komentar diprediksi sebagai negatif dan 7 sebagai netral.



Gambar 3. Confusion matrix Model Hybrid IndoBERT–IndoBERTweet

Berdasarkan Gambar 3, kelas positif memiliki jumlah prediksi benar terbanyak, yaitu 128 komentar. Sementara itu, kelas netral memiliki proporsi kesalahan paling tinggi dibandingkan kelas lainnya. Pada dataset penelitian ini, komentar netral cenderung sulit dibedakan karena dapat berbentuk informasi atau pertanyaan yang tetap mengandung kata-kata bernuansa positif maupun negatif.

Tabel 8. Contoh Kesalahan Prediksi Kelas Netral

Komentar	Label Aktual	Prediksi	Analisis linguistik
“Tidak mengurangi beban rupiah, tetapi alasannya cukup masuk akal.”	Netral	Positif	Komentar mengandung negasi dan <i>mixed sentiment</i> sehingga model lebih menangkap bagian positif.
“untuk redenominasi sepertinya akan sulit di realisasikan karena	Netral	Negatif	Komentar mengandung <i>mixed sentiment</i> . Kata “sulit”,

redenominasi sebaiknya dilakukan saat ekonomi stabil tidak harus sangat kuat tapi tidak dianjurkan ketika ekonomi sedang lemah atau inflasi tinggi tapi karena harapan target tercapai di tahun seharusnya ekonomi sudah cukup bagus untuk redenominasi”			“ekonomi lemah”, dan “inflasi tinggi” memberi kecenderungan negatif, sedangkan frasa “harapan target tercapai” dan “ekonomi sudah cukup bagus” menunjukkan optimisme. Penggunaan kata penghubung “tapi” memperlihatkan peralihan dari kekhawatiran menuju harapan, sehingga sentimen keseluruhan cenderung netral.
--	--	--	--

3.7. Pembahasan

Berdasarkan pengujian pada *seed* 42, model *hybrid* IndoBERT–IndoBERTtweet memperoleh nilai akurasi 76,73% dan Macro F1-score 74,83%, yang merupakan hasil tertinggi pada eksperimen ini. Berdasarkan akurasi, model *hybrid* diikuti oleh IndoBERTtweet, IndoBERT, dan SVM, sedangkan berdasarkan Macro F1-score, model *hybrid* diikuti oleh IndoBERT, IndoBERTtweet, dan SVM. Pengujian pada *seed* 42, 52, dan 62 juga menunjukkan bahwa model *hybrid* memiliki nilai rata-rata akurasi dan Macro F1-score tertinggi pada dataset dan konfigurasi yang digunakan. Hasil tersebut mengindikasikan bahwa penggabungan probabilitas IndoBERT dan IndoBERTtweet melalui *weighted soft voting* berpotensi meningkatkan kinerja klasifikasi pada dataset penelitian ini dibandingkan model tunggal. Namun, hasil ini belum dapat digeneralisasikan karena penelitian hanya menggunakan komentar dari satu video YouTube, konfigurasi model tertentu, dan jumlah *seed* yang terbatas. Pada model *hybrid*, kelas netral masih memiliki tingkat kesalahan prediksi yang relatif tinggi. Analisis komentar menunjukkan bahwa pertanyaan, negasi, dan *mixed sentiment* dapat menyebabkan komentar netral diprediksi sebagai positif atau negatif.

4. KESIMPULAN

Penelitian ini membandingkan kinerja SVM, IndoBERT, IndoBERTtweet, dan *hybrid* IndoBERT–IndoBERTtweet dalam mengklasifikasikan sentimen komentar YouTube mengenai redenominasi rupiah. Pada *seed* 42, model *hybrid* memperoleh kinerja tertinggi dengan akurasi 76,73% dan Macro F1-score 74,83%, sedangkan pengujian pada *seed* 42, 52, dan 62 menghasilkan rata-rata akurasi $76,00\% \pm 0,65\%$ dan Macro F1-score $73,74\% \pm 0,96\%$. Hasil ini menunjukkan bahwa pada dataset dan konfigurasi penelitian ini, metode *weighted soft voting* cenderung memberikan kinerja lebih baik dibandingkan model tunggal. Kelas netral masih relatif sulit diklasifikasikan karena adanya pertanyaan, negasi, dan *mixed sentiment*. Penelitian ini terbatas pada 1.588 komentar dari satu video YouTube, distribusi kelas yang tidak sepenuhnya seimbang, penggunaan satu pembagian data, serta belum adanya pengujian lintas waktu dan topik. Oleh karena itu, penelitian selanjutnya disarankan menggunakan data yang lebih besar dan beragam dari beberapa video atau kanal, menerapkan beberapa pembagian data atau validasi silang, serta menguji metode *hybrid* lain seperti *stacking* dan *majority voting*.

REFERENSI

- [1] N. B. Soleh and S. Herianingrum, “Redenomination: A Critical Review of Islamic Economics Based on Maqashid Sharia,” *Airlangga J. Innov. Manag.*, vol. 6, no. 2, pp. 378–393, 2025, doi: 10.20473/ajim.v6i2.72671.
- [2] M. Arsyad, “Dampak Redenominasi Rupiah terhadap Penyajian Laporan Keuangan Audited,” *AKUNSIKA J. Akunt. dan Keuang.*, vol. 4, no. 2, pp. 57–62, 2023, doi: 10.31963/akunsika.v4i2.4277.

- [3] B. P. Aulia, K. A. Pamungkas, L. M. Mauboy, and T. K. Wijayanti, "Forecasting Money Supply Menggunakan Metode GARCH untuk Kebijakan Redenominasi Indonesia," *J. BPPK*, vol. 17, no. 2, pp. 37–50, 2024.
- [4] R. Rizaldi and Y. Zamaya, "Perception of Micro, Small and Medium Enterprises (MSMEs) in Pekanbaru towards the Rupiah Redenomination Policy Plan," *Sosio e-Kons*, vol. 17, no. 1, pp. 30–41, 2025, doi: 10.30998/sosioekons.v17i1.27766.
- [5] A. Rustanta, S. Dwi Putranto, and P. Huang, "Maintaining the Digital Public Space: Communication Ethics and Regulatory Challenges in the TikTok Era," *J. Komun.*, vol. 17, no. 1, pp. 63–83, 2025, doi: 10.24912/jk.v17i1.32927.
- [6] B. Mulyono, I. Affandi, K. Suryadi, and C. Darmawan, "Online civic engagement through social media: An analysis of Twitter's big data," *Cakrawala Pendidik.*, vol. 42, no. 1, pp. 12–26, 2023, doi: 10.21831/cp.v42i1.54201.
- [7] B. Wicaksono and V. R. S. Nastiti, "Analisis Sentimen dalam Opini Publik di Chanel Youtube Indonesia Lawyers Club Tentang Isu Populer dengan Menggunakan Metode LSTM dan Bi-LSTM," *J. Algoritma.*, vol. 21, no. 2, pp. 241–251, 2024, doi: 10.33364/algoritma/v.21-2.1696.
- [8] A. N. A. Saputra, R. E. Saputro, and D. I. S. Saputra, "Enhancing Sentiment Analysis Accuracy Using SVM and Slang Word Normalization on YouTube Comments," *Sink. J. dan Penelit. Tek. Inform.*, vol. 9, no. 2, pp. 687–699, 2025, doi: 10.33395/sinkron.v9i2.14613.
- [9] O. I. Gifari, M. Adha, I. R. Hendrawan, and F. F. S. Durrand, "Analisis Sentimen Review Film Menggunakan TF-IDF dan Support Vector Machine," *J. Inf. Technol.*, vol. 2, no. 1, pp. 36–40, 2022, doi: 10.46229/jifotech.v2i1.330.
- [10] J. E. Br Sinulingga and H. C. K. Sitorus, "Analisis Sentimen Opini Masyarakat terhadap Film Horor Indonesia Menggunakan Metode SVM dan TF-IDF," *J. Manaj. Inform.*, vol. 14, no. 1, pp. 42–53, 2024, doi: 10.34010/jamika.v14i1.11946.
- [11] V. B. Lestari and C. A. Hutagalung, "Evaluation of TF-IDF Extraction Techniques in Sentiment Analysis of Indonesian-Language Marketplaces Using SVM, Logistic Regression, and Naive Bayes," *J. Comput. Sci. Appl.*, vol. 8, no. 1, pp. 36–44, 2025, doi: 10.21009/j-koma.v8i1.05.
- [12] N. N. A. Aryanti and O. Suria, "Analisis Sentimen terhadap Pemutusan Hubungan Kerja di Indonesia: Komparasi IndoBERT dengan SVM, Random Forest, dan Decision Tree dengan Optimasi TF-IDF," *RABITJ. Teknol. dan Sist. Inf. Univrab*, vol. 10, no. 2, pp. 1158–1176, 2025, doi: 10.36341/rabit.v10i2.6364.
- [13] W. V. Meilasroh, A. E. Pajri, and D. Nurdiansyah, "Analisis Sentimen Publik Terhadap ChatGPT: Perbandingan Logistic Regression Dan Distilbert," *Decod. J. Pendidik. Teknol. Inf.*, vol. 6, no. 2, pp. 308–320, 2026, doi: 10.51454/decode.v6i2.1644.
- [14] M. F. Kono, I. N. Fajri, and Y. Pristyanto, "Public Sentiment Analysis on Corruption Issues in Indonesia Using IndoBERT Fine-Tuning, Logistic Regression, and Linear SVM," *J. Appl. Informatics Comput.*, vol. 9, no. 5, pp. 2616–2628, 2025, doi: 10.30871/jaic.v9i5.10537.
- [15] R. Nugraha and S. Febriani, "Analisis Sentimen Topik #Kaburajadulu di Platform X Berbasis Model IndoBERTweet," *J. Mandalika Lit.*, vol. 6, no. 3, pp. 1084–1095, 2025, doi: 10.36312/jml.v6i3.5338.
- [16] V. D. Setiawan, D. U. Iswavigra, and E. Anggiratih, "Implementation of IndoBERT for Sentiment Analysis of the Constitutional Court's Decision Regarding the Minimum Age of Vice Presidential Candidates," *Sci. J. Informatics*, vol. 12, no. 3, pp. 397–406, 2025, doi: 10.15294/sji.v12i3.26320.
- [17] N. Maulida, N. Suarna, and W. Prihartono, "Analisis Ulasan Sentimen Aplikasi Mobile JKN Dengan Algoritma Support Vector Machine Berbasis Particle Swarm Optimization," *J. Mhs. Tek. Inform.*, vol. 8, no. 2, pp. 1651–1658, 2024, doi: 10.36040/jati.v8i2.9105.
- [18] M. I. Sa'ad, L. Muflikhah, and F. A. Bachtiar, "Perbandingan SVM dan IndoBERT untuk

- Analisis Sentimen Layanan Akademik Mahasiswa,” *Bull. Inf. Technol.*, vol. 7, no. 2, pp. 265–277, 2026, doi: 10.47065/bit.v7i2.2913.
- [19] A. S. N. Chairat, R. Rizal, and H. Himawan, “Comparative Sentiment Analysis of YouTube Comments on Indonesia’s Electric Vehicle Incentive Policy Using TF-IDF and IndoBERTweet Models,” *J. Tek. Inform.*, vol. 6, no. 6, pp. 5988–6002, 2025, doi: 10.52436/1.jutif.2025.6.6.5499.
- [20] U. I. Shabrina, M. I. Java, and S. Rochimah, “Optimizing Sentiment Analysis in Educational YouTube Videos: A Comparative Study of RoBERTa and Multinomial Naïve Bayes,” *JUTI J. Ilm. Teknol. Inf.*, vol. 22, no. 2, pp. 83–90, 2024, doi: 10.12962/j24068535.v22i2.a1204.
- [21] N. C. Majid and A. D. Indriyanti, “Analisis Sentimen Terhadap RUU TNI Di Platform X (Twitter) Menggunakan Metode Ensemble Berbasis Naïve Bayes Dan Support Vector Machine,” *J. Informatics Comput. Sci.*, vol. 7, no. 2, pp. 426–434, 2025.