

# ANALISIS SENTIMEN PUBLIK TENTANG INVESTASI EMAS DAN BITCOIN MENGGUNAKAN INDOBERT

Oleh :  
Rachmad Aditya Alfaris

Dosen Pembimbing:  
Dr. Mochamad Alfian Rosid S.Kom., M.Kom

Program Studi Informatika  
Universitas Muhammadiyah Sidoarjo  
2025

# Latar Belakang

Perubahan global menuntut peningkatan literasi finansial masyarakat. Melalui media sosial, edukasi investasi emas dan Bitcoin semakin populer, tetapi sering menimbulkan perdebatan. Kajian ini menggunakan NLP berbasis IndoBERT untuk memahami persepsi publik di Indonesia.

# Rumusan Masalah

- ✧ Bagaimana membangun model deep learning berbasis arsitektur Transformers untuk analisis sentimen?
- ✧ Bagaimana kinerja model IndoBERT dalam mengklasifikasikan sentimen publik terhadap topik investasi seperti Emas dan Bitcoin dalam konteks Bahasa Indonesia?

# Tujuan

- \* Membangun model deep learning berbasis arsitektur Transformers untuk analisis sentiment.
- \* Menilai kinerja model IndoBERT dalam mengklasifikasikan sentimen publik terhadap topik investasi seperti Emas dan Bitcoin dalam konteks Bahasa Indonesia.

# Batasan Masalah

- ✧ **Sumber data:** komentar publik YouTube & Tiktok (kanal “Ngomongin Uang”).
- ✧ **Objek:** fokus dua topik yaitu Emas dan Bitcoin.
- ✧ **Cakupan:** klasifikasi & perbandingan sentimen (positif/negatif/netral); **bukan** prediksi harga.
- ✧ **Pengelolaan:** melalui deteksi & translasi, case folding, cleansing, normalisasi, tokenisasi.

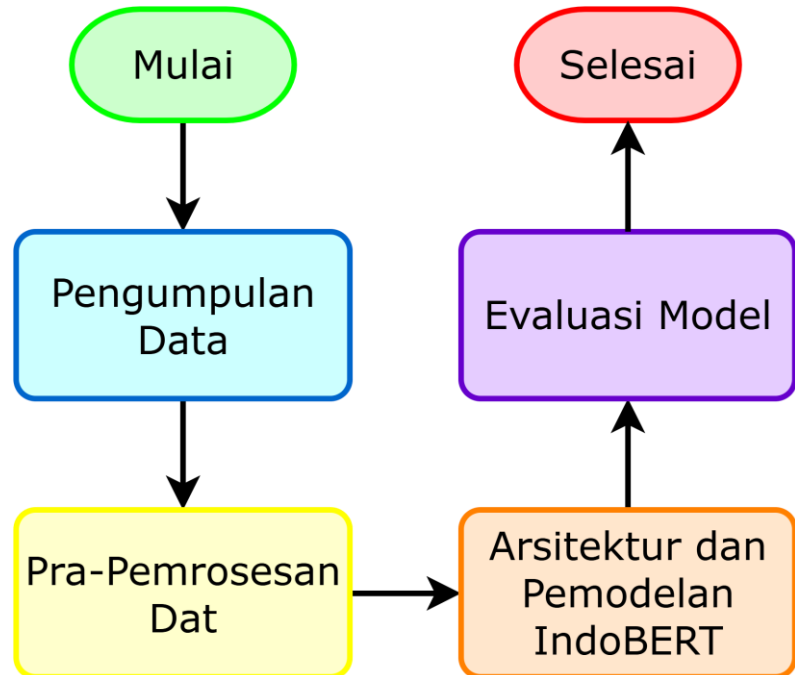
# Tinjauan Pustaka

- Sentiment Analysis on Twitter Social Media towards Climate Change on Indonesia Using IndoBERT Model (*Mubaraq & Maharani, 2022*)
- Analyzing the Indonesian Sentiment to Rohingya Refugees using IndoBERT Model (Arifin et al., 2024)
- Improving Sentiment Analysis and Topic Extraction in Indonesian Travel App Reviews Through BERT Fine-Tuning (*Irmawan et al., 2024*)
- Fine-Tuning IndoBERT untuk Analisis Sentimen pada Ulasan Pengguna Aplikasi Tiket.com di Google Play Store (*Nuryadi et al., 2025*)

# Analisis GAP

Studi tentang IndoBERT banyak di layanan publik/edukasi/aplikasi, tetapi **belum spesifik** membahas persepsi finansial **emas vs bitcoin** dari kanal edukasi di media sosial; penelitian ini mengisi celah tersebut.

# Alur Metodologi



1. Pengumpulan Data
2. Pra-Pemrosesan Data
3. Arsitektur & Pemodelan IndoBERT
4. Evaluasi Model

# Pengumpulan Data

- **YouTube:** YouTube Data API v3, kata kunci terkait “investasi emas”, “bitcoin untuk pemula”, dll.
- **TikTok:** **Apify** untuk crawling komentar bertopik emas/Bitcoin.
- Simpanan data: **CSV & XLSX**.

# Pra-Pemrosesan Data

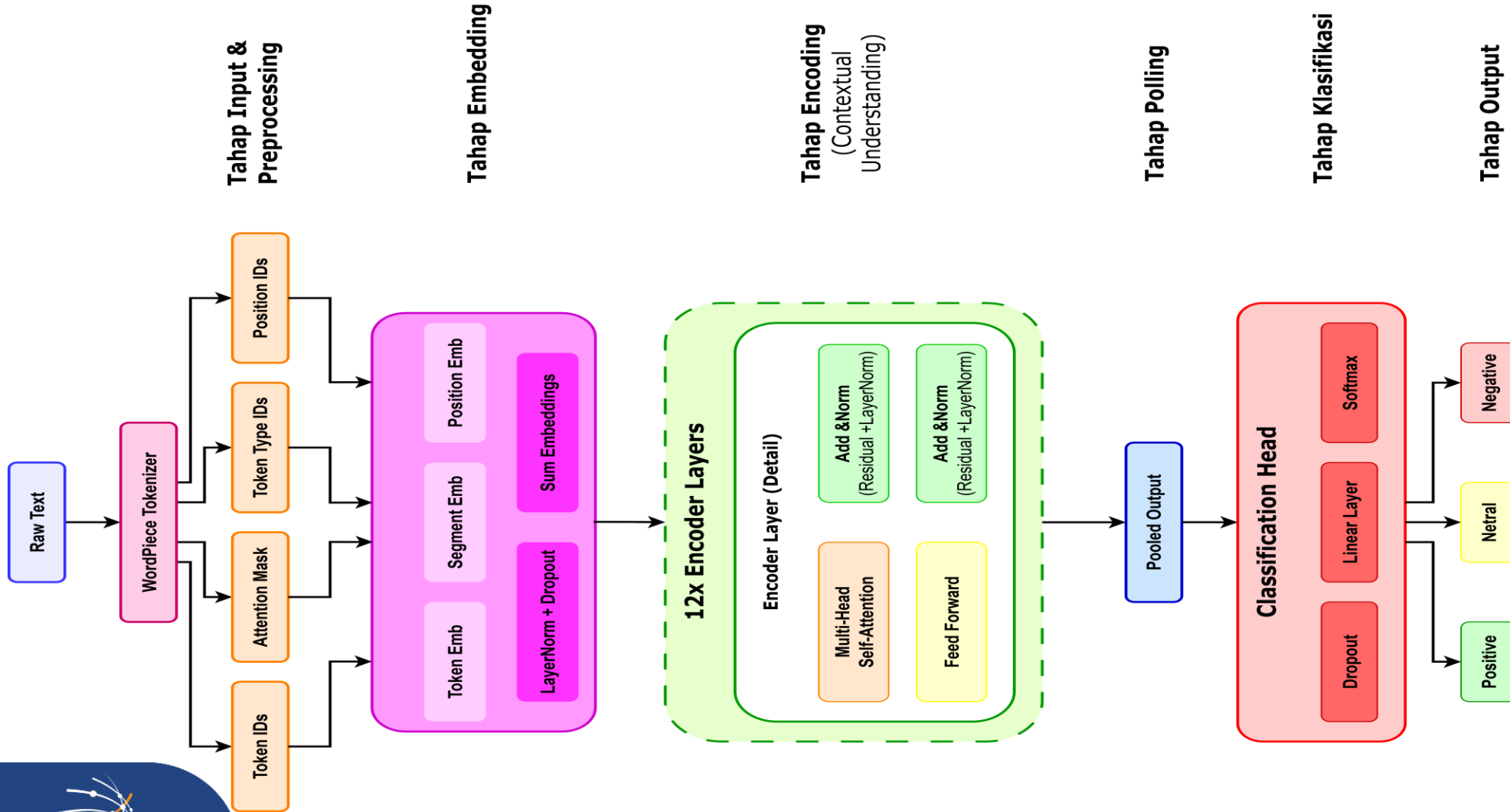
1. Deteksi & Translasi Bahasa (EN → ID bila perlu)
2. Case Folding
3. Cleansing (hapus emoji, angka, tautan, tanda baca tertentu, karakter khusus)
4. Normalisasi (kata tidak baku/typo → baku)
5. Tokenisasi

Setelah pra-pemrosesan, komentar diberi label **2=Positif**, **1=Netral**, **0=Negatif**, lalu dibagi **train 80%** / **test 10%** / **validation 10%** sebelum pemodelan.

# Arsitektur dan Pemodelan IndoBERT

IndoBERT adalah model BERT berbahasa Indonesia (bidirectional context) yang efektif untuk tugas NLP seperti analisis sentimen; relevan karena komentar media sosial bersifat informal dan kontekstual.

# Arsitektur dan Pemodelan IndoBERT



# Evaluasi Model

- **Accuracy** — proporsi prediksi benar.
- **Precision** — ketepatan prediksi per kelas.
- **Recall** — cakupan deteksi instance per kelas.
- **F1-Score** — rata-rata harmonis precision & recall.
- **Confusion Matrix** — distribusi prediksi vs label aktual ( $3 \times 3$ ).

# Karakteristik Dataset

- **Total Data Bersih:** 4.909 komentar (setelah pra-pemrosesan dan validasi).
- **Distribusi Topik:** Didominasi oleh Bitcoin (75,2% / 3.690 komentar) dibandingkan Emas (24,8% / 1.219 komentar).
- **Distribusi Sentimen:** Positif (45%), Netral (40,6%), dan Negatif (14,4%). Sentimen negatif merupakan kelas minoritas.
- **Penanganan Ketimpangan Kelas (Class Imbalance):** Menggunakan teknik *Class Weight*. Bobot kelas Negatif diberi penalti terbesar (2,3086), Netral (0,8207), dan Positif (0,7416).

# Skenario Eksperimen Fine-Tuning

- **Skenario Pelatihan:** Dibagi menjadi Skenario A (5 epoch) dan Skenario B (20–50 epoch) dengan total 10 konfigurasi eksperimen.
- **Model Terbaik (Best Model):** Konfigurasi **A4** terbukti paling optimal.  
Parameter **A4**: 5 *epochs*, menggunakan *Class Weight*, penyisipan topik (*Prepend*), dan *Optimizer AdamW*.
- **Pencapaian Model A4:**
  - a. Akurasi Keseluruhan: **72,51%**
  - b. Test Macro F1-Score: **0,6983**
  - c. Waktu Pelatihan: Sangat efisien (hanya **2,4 menit**).

# Analisis Performa Model

- **Efek Prepend & Class Weight:** Kombinasi kedua strategi ini sukses meningkatkan kemampuan model dalam mengenali kelas minoritas (F1-Score sentimen Negatif naik signifikan dari 0,5694 menjadi **0,6087**).
- **Fenomena Overfitting:**
  - a. Penambahan durasi pelatihan pada Skenario B (hingga 50 epoch) tidak meningkatkan performa.
  - b. Justru memicu *overfitting*, ditandai dengan *validation loss* yang melonjak tajam (mencapai angka 3,7).

# Stabilitas Model Lintas Topik

- **Tujuan Uji Stabilitas:** Memastikan model adil dan tidak bias terhadap salah satu instrumen investasi.
- **Hasil Evaluasi per Topik:**
  - a. Akurasi prediksi topik **Bitcoin (BTC): 72,70%**
  - b. Akurasi prediksi topik **Emas (XAU): 71,90%**
- **Kesimpulan Uji:** Selisih akurasi sangat kecil ( $< 1\%$ ), membuktikan model A4 memiliki stabilitas generalisasi yang sangat baik pada lintas topik literasi finansial.

# Kesimpulan

- Model IndoBERT varian *indobenchmark/indobert-base-p2* terbukti andal memetakan sentimen publik di sektor investasi finansial berbahasa Indonesia yang informal.
- Sinergi antara penyisipan topik (*prepend*) dan penyesuaian bobot (*class weight*) sangat krusial untuk mengatasi ketidakseimbangan data sentimen di media sosial.
- Pelatihan model (*fine-tuning*) yang efisien (siklus pendek/5 epoch) jauh lebih optimal dibandingkan siklus panjang yang justru memicu *overfitting*.

# Implikasi dan Saran

Studi ini diharapkan memberi masukan berbasis data bagi kreator konten dan pemangku kepentingan untuk merancang strategi komunikasi finansial yang lebih tepat serta meminimalkan miskonsepsi publik.

# Daftar Pustaka

- I. R. Johan and S. A. Azariani, "Persepsi Risiko, Financial Self-Efficacy dan Minat Investasi Emas pada Generasi Z.," *J. Ilmu Kel. dan Konsum.*, vol. 18, no. 1, pp. 26-37, 2025, doi: 10.24156/jikk.2025.18.1.26.
- D. N. Ahsanah, "Emas sebagai Instrumen Investasi Jangka Panjang," *Shar-E J. Kaji. Ekon. Huk. Syariah*, vol. 8, no. 1, pp. 177-187, 2022.
- A. Wicaksono, C. E. Violita, and E. R. Kamila, "Bitcoin sebagai Instrumen Investasi yang Menguntungkan," *Greenomika*, vol. 4, no. 1, p. 44, 2022.
- M. Hougan and D. Lawant, *Cryptoassets: The Guide to Bitcoin, Blockchain, and Cryptocurrency for Investment Professionals*. CFA Institute Research Foundation, 2021. [Online]. Available: <https://www.cfainstitute.org/research/foundation>
- D. Anisa, T. Anggraini, and K. Tambunan, "Analisis Cryptocurrency Sebagai Alat Alternatif Berinvestasi di Indonesia," *Own. Ris. dan J. Akunt.*, vol. 7, no. 3, 2023, doi: 10.33395/owner.v7i3.1698.
- M. A. Rosid, D. O. Siahaan, and A. Saikhu, "Sarcasm Detection in Indonesian-English Code-Mixed Text Using Multihead Attention-Based Convolutional and Bi-Directional GRU," *IEEE Access*, 2024, doi: 10.1109/ACCESS.2024.3436107.
- R. I. Perwira, V. A. Permadi, D. I. Purnamasari, and R. P. Agusdin, "Domain-Specific Fine-Tuning of IndoBERT for Aspect-Based Sentiment Analysis in Indonesian Travel User-Generated Content," *J. Inf. Syst. Eng. Bus. Intell.*, vol. 11, no. 1, pp. 30-40, 2025, doi: 10.20473/jisebi.11.1.30-40.

N. Alfarobby and H. Irawan, "Analisis Sentimen Kepuasan Konsumen Pengguna Transportasi Online pada Ulasan Google Playstore Menggunakan Indobert dan Topic Modeling (Studi Kasus: Gojek dan Grab)," *e-Proceeding Manag.*, vol. 11, no. 1, p. 72, 2024.

S. Riyadi, L. K. Salsabila, C. Damarjati, and R. Abdul Karim, "Sentiment Analysis of YouTube Users on Blackpink Kpop Group Using IndoBERT," *INTENSIF J. Ilm. Penelit. dan Penerapan Teknol. Sist. Inf.*, vol. 8, no. 2, p. 233, 2024, doi: 10.29407/intensif.v8n2.22678.

S. Aras, M. Yusuf, R. Ruimassa, E. A. B. Wambrau, and E. B. Palalangan, "Sentiment Analysis on Shopee Product Reviews Using IndoBERT," *J. Inf. Syst. Informatics*, vol. 6, no. 3, p. 1616, 2024, doi: 10.51519/journalisi.v6i3.814.

W. M. Baihaqi and A. Munandar, "Sentiment Analysis of Student Comment on the College Performance Evaluation Questionnaire Using Naïve Bayes and IndoBERT," *JUITA J. Inform.*, vol. 11, no. 2, pp. 213-220, 2023.

Nurhasiyah, R. Dwiyanaputra, S. I. Murpratiwi, and A. Aranta, "Analisis Sentimen Pengguna Platform Media Sosial X pada Topik Pemilihan Presiden 2024 Menggunakan Perbandingan Model Monolingual dan Multilingual BERT," *JATI J. Mhs. Tek. Inform.*, vol. 9, no. 1, p. 626, 2025.

Jazuli, Widowati, and R. Kusumaningrum, "Aspect-Based Sentiment Analysis on Student Reviews Using the Indo-Bert Base Model," *E3S Web Conf.*, vol. 448, 2023, doi: 10.1051/e3sconf/202344802004.

C. Apriyadi and Styawati, "Sentiment Analysis of Cyber Attacks in Bank Syariah Indonesia Using SVM and Indobert Method," *J. Tek. Inform.*, vol. 6, no. 2, pp. 819-838, 2025, doi: 10.52436/1.jutif.2025.6.2.2636.

O. A. Irmawan, I. Budi, A. B. Santoso, and P. K. Putra, "Improving Sentiment Analysis and Topic Extraction in Indonesian Travel App Reviews Through BERT Fine-Tuning," *J. Nas. Pendidik. Tek. Inform.*, vol. 13, no. 2, p. 359, 2024.



# Daftar Pustaka

- Angelica, A. E., & Zen, T. S. (2023). Effect of Financial Contents on Social Media Towards Financial Literacy on Generation Z in Sumatra and Java. *Journal of Consumer Studies and Applied Marketing*, 1(1), 37–45. <https://doi.org/10.58229/jcsam.v1i1.47>
- Apriyadi, C., & Styawati, S. (2025a). Sentiment Analysis of Cyber Attacks in Bank Syariah Indonesia Using SVM and Indobert Method. *Jurnal Teknik Informatika (Jutif)*, 6(2), 819–838. <https://doi.org/10.52436/1.jutif.2025.6.2.2636>
- Apriyadi, C., & Styawati, S. (2025b). Sentiment Analysis of Cyber Attacks in Bank Syariah Indonesia Using SVM and Indobert Method. *Jurnal Teknik Informatika (Jutif)*, 6(2), 819–838. <https://doi.org/10.52436/1.jutif.2025.6.2.2636>
- Ayu, F., Aryanti, D., Luthfiarta, A., Adiwinata, D., & Soeroso, I. (2025). Aspect-Based Sentiment Analysis with LDA and IndoBERT Algorithm on Mental Health App: Riliv. In *Journal of Applied Informatics and Computing (JAIC)* (Vol. 9, Issue 2). <http://jurnal.polibatam.ac.id/index.php/JAIC>
- Febrianto, D. C., Fitriani, M. A., Afrad, M., & Khadija, M. A. (n.d.). *ASPECT BASED SENTIMENT ANALYSIS MENGGUNAKAN INDOBERT MODEL TERHADAP REVIEW PENGUNJUNG OBJEK WISATA BATURRADEN* (Vol. 10).
- Firdhy, E. H. (n.d.). *ANALISIS PERBANDINGAN KINERJA CRYPTOCURRENCY BITCOIN, SAHAM DAN EMAS SEBAGAI ALTERNATIF INVESTASI Lailatul Amanah Sekolah Tinggi Ilmu Ekonomi Indonesia (STIESIA) Surabaya*.
- Irmawan, O. A., Budi, I., Santoso, A. B., & Putra, P. K. (2024). Improving Sentiment Analysis and Topic Extraction in Indonesian Travel App Reviews Through BERT Fine-Tuning. *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, 13(2), 359–370. <https://doi.org/10.23887/janapati.v13i2.77028>
- Jayadianti, H., Kaswidjanti, W., Utomo, A. T., Saifullah, S., Dwiyanto, F. A., & Drezewski, R. (2022). Sentiment analysis of Indonesian reviews using fine-tuning IndoBERT and R-CNN. *ILKOM Jurnal Ilmiah*, 14(3), 348–354. <https://doi.org/10.33096/ilkom.v14i3.1505.348-354>

Maemunah, S. (2022). Strategies for enhancing financial literacy in Indonesia. *Journal of Economics and Business Letters*, 2(1), 13–15. <https://doi.org/10.55942/jebll.v2i1.283>

Mubaraq, M. F., & Maharani, W. (2022). Sentiment Analysis on Twitter Social Media towards Climate Change on Indonesia Using IndoBERT Model. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 6(4), 2426. <https://doi.org/10.30865/mib.v6i4.4570>

Munandar, W. M. (2023). *Sentiment Analysis of Student Comment on the College Performance Evaluation Questionnaire Using Naïve Bayes and IndoBERT* (Vol. 11, Issue 2). <https://github.com/notfound313/sentimen->

Nakamoto, S. (n.d.). *Bitcoin: A Peer-to-Peer Electronic Cash System*. [www.bitcoin.org](http://www.bitcoin.org)

Nuryadi, D., Metandi, F., Alam Hadiwijaya, N., Zainul Rohman, M., Hartanto, S., Syafrizal, A., Yadie Teknologi Informasi, E., Negeri Samarinda Jl Cipto Mangunkusumo, P., Seberang, S., & Timur, K. (2025). FINE TUNING INDOBERT UNTUK ANALISIS SENTIMEN PADA ULASAN PENGGUNA APLIKASI TIKET.COM DI GOOGLE PLAY STORE. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 9, Issue 2).

Riyadi, S., Salsabila, L. K., Damarjati, C., & Karim, R. A. (2024). INTENSIF: Jurnal Ilmiah Penelitian dan Penerapan Teknologi Sistem Informasi 233 Sentiment Analysis of YouTube Users on Blackpink Kpop Group Using IndoBERT 1 \*. *INTENSIF*, 8(2), 2549–6824. <https://doi.org/10.29407/intensif.v8n2.22678>

Rosid, M. A., Siahaan, D., & Saikhu, A. (2024). Sarcasm Detection in Indonesian-English Code-Mixed Text Using Multihead Attention-Based Convolutional and Bi-Directional GRU. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2024.3436107>

Setiawan, J. C., M. Lhaksmana, K., & Bunyamin, B. (2023). Sentiment Analysis of Indonesian TikTok Review Using LSTM and IndoBERTweet Algorithm. *JUPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, 8(3), 774–780. <https://doi.org/10.29100/jipi.v8i3.3911>