

202110370311169
Fathul Agit Darmawan
Prodi Informatika

**PENDEKATAN HIBRIDA PERINGKASAN TRANSKRIPSI
VIDEO DENGAN KLASTERISASI K-MEANS BERBASIS BERT
DAN BART**

TUGAS AKHIR

Diajukan Untuk Memenuhi
Persyaratan Guna Meraih Gelar Sarjana
Informatika Universitas Muhammadiyah Malang



Fathul Agit Darmawan

202110370311169

Data Science

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH MALANG**

2025

202110370311169
Fathul Agit Darmawan
Prodi Informatika

LEMBAR PERSETUJUAN

Pendekatan Hibrida Peringkasan Transkripsi Video dengan Klasterisasi K-Means Berbasis BERT dan BART

TUGAS AKHIR

**Sebagai Persyaratan Guna Meraih Gelar Sarjana Strata 1
Informatika Universitas Muhammadiyah Malang**

Menyetujui,
Malang, *16 April 2026*

Dosen Pembimbing 1



Christian Sri Kusuma Aditva S.Kom.,

M.Kom

NIP. 180327021991PNS.

Dosen Pembimbing 2



=

NIP.

LEMBAR PENGESAHAN

Pendekatan Hibrida Peringkasan Transkripsi Video dengan Klasterisasi K-Means Berbasis BERT dan BART

TUGAS AKHIR

Sebagai Persyaratan Guna Meraih Gelar Sarjana Strata 1
Informatika Universitas Muhammadiyah Malang

Disusun Oleh :

FATHUL AGIT DARMAWAN

202110370311169

Tugas Akhir ini telah diuji dan dinyatakan lulus melalui sidang majelis penguji
pada tanggal 16 April 2026

Menyetujui,

Dosen Pembimbing 1



Christian Sri Kusuma Aditya S.Kom.,

M.Kom

NIP. 180327021991PNS.

Dosen Pembimbing 2



=
NIP.

Dosen Penguji 1



Hardianto Wibowo S.Kom, MT.

NIP. 10816120592PNS.

Dosen Penguji 2



Sofyan Arifianto S.Si., M.Kom

NIP. 10818030646PNS.

Mengetahui,
Ketua Jurusan Informatika



Iri Agus Eko Minarno S.Kom., M.Kom. IPM.

NIP. 10814100540PNS.



LEMBAR PERNYATAAN

Yang bertanda tangan dibawah ini :

NAMA : FATHUL AGIT DARMAWAN

NIM : 202110370311169

FAK./JUR. : INFORMATIKA

Dengan ini saya menyatakan bahwa Tugas Akhir dengan judul "**Pendekatan Hibrida Peringkasan Transkripsi Video dengan Klasterisasi K-Means Berbasis BERT dan BART**" beserta seluruh isinya adalah karya saya sendiri dan bukan merupakan karya tulis orang lain, baik sebagian maupun seluruhnya, kecuali dalam bentuk kutipan yang telah disebutkan sumbernya.

Demikian surat pernyataan ini saya buat dengan sebenar-benarnya. Apabila kemudian ditemukan adanya pelanggaran terhadap etika keilmuan dalam karya saya ini, atau ada klaim dari pihak lain terhadap keaslian karya saya ini maka saya siap menanggung segala bentuk resiko/sanksi yang berlaku.

Malang, 20 Agustus 2025

Mengetahui,
Dosen Pembimbing

Yang membuat pernyataan



Christian Sri Kusuma Aditya, S.Kom., M.Kom.

Fathul Agit Darmawan

ABSTRAK

Penggunaan video sebagai sarana penyampaian informasi dan edukasi terus meningkat di berbagai platform daring. Namun, durasi yang panjang serta penyampaian yang tidak terstruktur sering kali membuat penonton kesulitan memahami inti materi, sehingga menimbulkan tantangan dalam pengembangan metode peringkasan otomatis untuk monolog, wawancara, dan podcast. Metode ekstraktif cenderung menghasilkan ringkasan yang kurang mengalir, sedangkan metode abstraktif berisiko menghilangkan informasi penting. Untuk menjawab permasalahan tersebut, penelitian ini mengusulkan pendekatan hibrida yang menggabungkan teknik ekstraktif dan abstraktif. Pada tahap ekstraktif, kalimat direpresentasikan menggunakan embedding BERT dan dikelompokkan menggunakan metode K-Means Clustering. Tahap abstraktif kemudian memanfaatkan model BART untuk menghasilkan ringkasan yang lebih koheren dan informatif. Hasil evaluasi pada 20 video Human Metapneumovirus (HMPV) menunjukkan performa terbaik pada jenis monolog, dengan ROUGE-1 sebesar 57%, ROUGE-2 sebesar 30%, dan ROUGE-L sebesar 32%. Meskipun wawancara dan podcast menunjukkan hasil lebih rendah karena interaksi pembicara yang dinamis, pendekatan hibrida ini tetap unggul dibandingkan metode ekstraktif saja maupun abstraktif saja. Temuan ini menegaskan potensi pendekatan hibrida dalam menghasilkan peringkasan video yang lebih adaptif di masa mendatang.

Kata Kunci ~ Peringkasan Video; Pendekatan Hibrida; BERT; Clustering; K-Means; BART

ABSTRACT

The use of video as a medium for delivering information and education continues to grow across online platforms. However, long durations and unstructured delivery often make it difficult for viewers to grasp the main points, creating challenges for developing automatic summarization methods for monologues, interviews, and podcasts. Extractive methods tend to produce less cohesive summaries, while abstractive methods risk omitting important details. To address these issues, this study proposes a hybrid approach that integrates extractive and abstractive techniques. In the extractive stage, sentences are represented using BERT embeddings and grouped using the K-Means Clustering method. The abstractive stage then employs the BART model to generate summaries that are more coherent and informative. Experimental evaluations on 20 Human Metapneumovirus (HMPV) videos show the strongest performance for monologue-type content, achieving ROUGE-1 of 57%, ROUGE-2 of 30%, and ROUGE-L of 32%. Although lower performance was observed for interviews and podcasts due to dynamic interactions and frequent speaker shifts, the hybrid approach consistently outperformed both extractive-only and abstractive-only baselines. These findings highlight the potential of the hybrid approach in producing more adaptive video summarization in the future.

Keywords ~ Video Summarization; Hybrid Approach; BERT; Clustering; K-Means; BART

LEMBAR PERSEMBAHAN

Segala puji dan syukur saya panjatkan kepada Allah SWT atas segala rahmat, karunia, dan kekuatan yang telah diberikan sehingga karya ini dapat terselesaikan dengan baik. Dengan penuh kerendahan hati, karya ini saya persembahkan kepada :

1. **Keluarga tercinta**, yang selalu menjadi sumber kekuatan dan inspirasi dalam setiap langkah kehidupan saya. Terima kasih atas doa yang tiada henti, kasih sayang yang tulus, serta dukungan yang tidak pernah surut. Kalian adalah alasan utama saya untuk terus berjuang dan berusaha memberikan yang terbaik.
2. **Diri saya sendiri**, sebagai bentuk penghargaan atas usaha, kerja keras, pengorbanan, dan kesabaran dalam menghadapi berbagai rintangan selama proses ini. Karya ini menjadi bukti bahwa setiap usaha yang konsisten dan penuh keyakinan akan membuahkan hasil.
3. **Bapak Christian Sri Kusuma Aditya, S.Kom., M.Kom.**, selaku dosen pembimbing yang telah dengan sabar memberikan bimbingan, arahan, serta masukan yang sangat berarti dalam penyusunan karya ini. Terima kasih atas ilmu, perhatian, dan dukungan yang telah diberikan sehingga penelitian ini dapat terselesaikan dengan baik.
4. **Teman-teman terdekat**, yang senantiasa memberikan semangat, dukungan moral, serta menemani saya dalam suka maupun duka selama proses penyelesaian karya ini. Terima kasih atas kebersamaan, canda tawa, dan motivasi yang membuat perjalanan ini terasa lebih ringan dan bermakna.
5. **Almamater Universitas Muhammadiyah Malang**, serta **seluruh pihak yang telah membantu, mendukung, dan berkontribusi, baik secara langsung maupun tidak langsung**, yang tidak dapat saya sebutkan satu per satu. Semoga segala bantuan, dukungan, dan doa yang diberikan menjadi amal kebaikan yang senantiasa dibalas oleh Allah SWT.

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT atas rahmat dan hidayah-Nya sehingga penelitian ini dapat diselesaikan dengan baik. Karya ini disusun sebagai bagian dari proses akademik di Universitas Muhammadiyah Malang dengan judul **“Peringkasan Ekstraktif untuk Transkripsi Video Menggunakan Representasi BERT dan Algoritma K-Means.”**

Penelitian ini berfokus pada tantangan dalam mengolah transkripsi video yang panjang, repetitif, dan tidak terstruktur. Metode peringkasan ekstraktif berbasis BERT dikombinasikan dengan algoritma K-Means digunakan untuk menghasilkan ringkasan yang lebih singkat, padat, dan representatif. Evaluasi kualitas ringkasan dilakukan menggunakan metrik ROUGE sebagai ukuran perbandingan dengan ringkasan manual.

Penulis menyampaikan terima kasih kepada Bapak Christian Sri Kusuma Aditya, S.Kom., M.Kom. selaku dosen pembimbing atas bimbingan dan arahnya, kepada keluarga tercinta atas doa dan dukungannya, serta kepada teman-teman dan seluruh pihak yang telah membantu. Penulis menyadari karya ini masih jauh dari sempurna, sehingga kritik dan saran yang membangun sangat diharapkan. Semoga penelitian ini bermanfaat dalam pengembangan ilmu, khususnya di bidang pemrosesan bahasa alami dan peringkasan teks otomatis.

Malang, 20 Agustus 2025



Fathul Agit Darmawan

DAFTAR ISI

JUDUL	I
LEMBAR PERSETUJUAN	II
LEMBAR PENGESAHAN	III
LEMBAR PERNYATAAN	IV
ABSTRAK	V
ABSTRACT	VI
LEMBAR PERSEMBAHAN	VII
KATA PENGANTAR	VIII
DAFTAR ISI	IX
DAFTAR TABEL	XI
DAFTAR GAMBAR	XII
DAFTAR RUMUS	XIII
BAB I PENDAHULUAN	1
1.1. LATAR BELAKANG	1
1.2. RUMUSAN MASALAH.....	3
1.3. TUJUAN PENELITIAN.....	3
1.4. BATASAN PENELITIAN	4
BAB II TINJAUAN PUSTAKA	5
2.1. PERINGKASAN TEKS (<i>TEXT SUMMARIZATION</i>).....	5
2.2. TRANSKRIPSI VIDEO.....	5
2.3. MODEL BERT UNTUK REPRESENTASI KALIMAT.....	6
2.4. ALGORITMA K-MEANS CLUSTERING	7
2.5. MODEL BART UNTUK PERINGKASAN ABSTRAKTIF	7
2.6. METRIK EVALUASI ROUGE.....	9
BAB III METODOLOGI PENELITIAN	10
3.1. RANCANGAN PENELITIAN	10
3.2. SUMBER DAN AKUISISI DATA	11
3.3. PROSEDUR PENELITIAN	13
BAB IV HASIL DAN PEMBAHASAN	18
4.1. HASIL TRANSKRIPSI.....	18

202110370311169
Fathul Agit Darmawan
Prodi Informatika

4.2.	HASIL RINGKASAN EKSTRAKTIF	19
4.3.	HASIL RINGKASAN ABSTRAKTIF	29
4.4.	EVALUASI KUALITAS RINGKASAN MENGGUNAKAN ROUGE	33
4.5.	PERBANDINGAN DENGAN PENELITIAN SEBELUMNYA.....	36
4.6.	SARAN UNTUK PENELITIAN SELANJUTNYA	37
BAB V KESIMPULAN		39
DAFTAR PUSTAKA		41

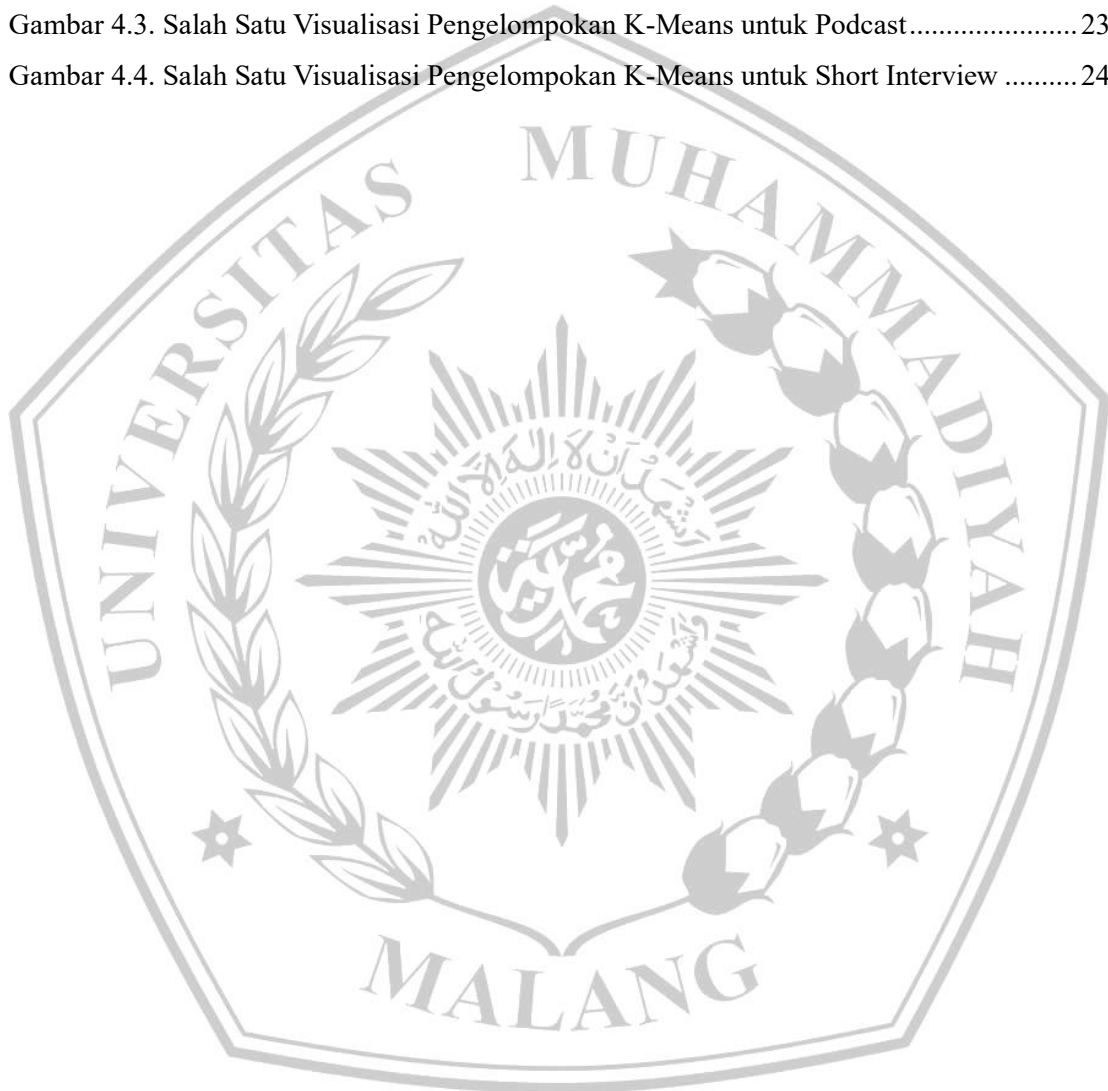


DAFTAR TABEL

Tabel 3.1. Detail Video yang Digunakan.....	11
Tabel 3.2. Konfigurasi Parameter BART untuk Proses Abstractive Summarization.....	16
Tabel 3.3. Menjelaskan komponen dari metrik ROUGE.....	17
Tabel 4.1. Hasil Transkripsi Video.....	18
Tabel 4.2. Pedoman Panjang Ringkasan Berdasarkan Jumlah Kalimat	19
Tabel 4.3. Salah Satu Contoh Representatif yang Dipilih dari Video Monolog.....	25
Tabel 4.4. Nilai Cluster dan Jumlah Klaster yang Dihasilkan per Video	26
Tabel 4.5. Jumlah Kata dan Kalimat dalam Hasil Ringkasan Ekstraktif.....	28
Tabel 4.6. Jumlah Chunk yang Dihasilkan per Video.....	29
Tabel 4.7. Jumlah Kata dan Kalimat Ringkasan Abstraktif Akhir.....	32
Tabel 4.8. ROUGE F1 Scores untuk Ringkasan Ekstraktif.....	33
Tabel 4.9. Perbandingan Rata-rata Skor ROUGE F1	35
Tabel 4.10. Perbandingan Hasil Ringkasan Hibrida dengan Gurusamy et al. (2023)	37

DAFTAR GAMBAR

Gambar 3.1. Alur Penelitian untuk Ringkasan Otomatis Video HMPV.....	10
Gambar 4.1. Salah Satu Visualisasi Pengelompokan K-Means untuk Monolog.....	21
Gambar 4.2. Salah Satu Visualisasi Pengelompokan K-Means untuk Interview	22
Gambar 4.3. Salah Satu Visualisasi Pengelompokan K-Means untuk Podcast.....	23
Gambar 4.4. Salah Satu Visualisasi Pengelompokan K-Means untuk Short Interview	24



DAFTAR RUMUS

Eq. 2.1	7
Eq. 2.2	8
Eq. 2.3	8
Eq. 2.4	9
Eq. 3.1	14
Eq. 3.2	15
Eq. 3.3	17



DAFTAR PUSTAKA

- [1] J. P. Suharsono and D. Nurahman, "Pemanfaatan Youtube Sebagai Media Peningkatan Pelayanan Dan Informasi," *Ganaya J. Ilmu Sos. Dan Hum.*, vol. 7, no. 1, pp. 298–304, Mar. 2024, doi: 10.37329/ganaya.v7i1.3157.
- [2] L. Parabhoi, R. R. Sahu, R. S. Dewey, M. K. Verma, A. Kumar Seth, and D. Parabhoi, "YouTube as a source of information during the Covid-19 pandemic: a content analysis of YouTube videos published during January to March 2020," *BMC Med. Inform. Decis. Mak.*, vol. 21, no. 1, p. 255, Dec. 2021, doi: 10.1186/s12911-021-01613-8.
- [3] A. Koenecke, A. S. G. Choi, K. X. Mei, H. Schellmann, and M. Sloane, "Careless Whisper: Speech-to-Text Hallucination Harms," 2024, doi: 10.48550/ARXIV.2402.08021.
- [4] J. K. Höhne, T. Lenzner, and J. Claassen, "Automatic speech-to-text transcription: evidence from a smartphone survey with voice answers," *Int. J. Soc. Res. Methodol.*, pp. 1–8, Jan. 2025, doi: 10.1080/13645579.2024.2443633.
- [5] Y. Qiu and Y. Jin, "Engineering Document Summarization: A Bidirectional Language Model-Based Approach," *J. Comput. Inf. Sci. Eng.*, vol. 22, no. 6, p. 061004, Dec. 2022, doi: 10.1115/1.4054203.
- [6] Le Ngoc Thang, Nguyen Minh Tien, Do Nhat Minh, Nguyen Chi Thanh, and Le Quang Minh, "A method to utilize prior knowledge for extractive summarization based on pre-trained language models," *Vietnam J. Sci. Technol.*, Dec. 2024, doi: 10.15625/2525-2518/20241.
- [7] Z. Edress and Y. Ortakci, "Optimizing Text Summarization with Sentence Clustering and Natural Language Processing," *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 10, 2024, doi: 10.14569/IJACSA.2024.01510115.
- [8] D. Jiang, S. Cao, and S. Yang, "Abstractive summarization of long texts based on BERT and sequence-to-sequence model," in *2021 2nd International Conference*

- on Information Science and Education (ICISE-IE)*, Chongqing, China: IEEE, Nov. 2021, pp. 460–466. doi: 10.1109/ICISE-IE53922.2021.00112.
- [9] K. Akiyama, A. Tamura, T. Ninomiya, and T. Kajiwara, “Hie-BART: Abstractive Summarization by Hierarchical BART,” *J. Nat. Lang. Process.*, vol. 29, no. 3, pp. 835–853, 2022, doi: 10.5715/jnlp.29.835.
- [10] P. S. Mankar and A. B. Manwar, “An Exploration of Extractive method and Abstractive Method of Text Summarization with Various Approaches, Techniques and Datasets,” *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 12, no. 11, pp. 422–430, Nov. 2024, doi: 10.22214/ijraset.2024.65100.
- [11] S. Anggraini, N. Hayatin, and G. I. Marthasari, “Peringkasan Multi Dokumen Berita Dengan Pemilihan Kalimat Utama Berbasis Algoritma Cluster Importance Dengan Mempertimbangkan Posisi Kalimat,” *J. Repos.*, vol. 2, no. 1, Jan. 2024, doi: 10.22219/repositor.v2i1.31529.
- [12] G. Zakawaly, N. Hayatin, and V. R. S. Nastiti, “Improviasi Algoritma Dijkstra Pada Peringkasan Teks Otomatis Untuk Artikel Politik,” *J. Repos.*, vol. 5, no. 2, Jan. 2024, doi: 10.22219/repositor.v5i2.32061.
- [13] A. L. Haz, E. D. Fajrianti, N. Funabiki, and S. Sukaridhoto, “A Study of Audio-to-Text Conversion Software Using Whispers Model,” in *2023 Sixth International Conference on Vocational Education and Electrical Engineering (ICVEE)*, Surabaya, Indonesia: IEEE, Oct. 2023, pp. 268–273. doi: 10.1109/ICVEE59738.2023.10348186.
- [14] R. Wang, Z. Xu, and F. X. Lin, “WhisperFlow: speech foundation models in real time,” 2024, *arXiv*. doi: 10.48550/ARXIV.2412.11272.
- [15] D. Ferdiansyah and C. S. K. Aditya, “Implementasi Automatic Speech Recognition Bacaan Al-Qur’an Menggunakan Metode Wav2Vec 2.0 dan OpenAI-Whisper,” *J. Tek. Elektro Dan Komput. TRIAC*, vol. 11, no. 1, pp. 11–16, Jun. 2024, doi: 10.21107/triac.v11i1.24332.

- [16] Z. Zhang *et al.*, “Semantics-Aware BERT for Language Understanding,” *Proc. AAAI Conf. Artif. Intell.*, vol. 34, no. 05, pp. 9628–9635, Apr. 2020, doi: 10.1609/aaai.v34i05.6510.
- [17] Z. Wei, X. Xu, C. Wang, Z. Liu, P. Xin, and W. Zhang, “An Index Construction and Similarity Retrieval Method Based on Sentence-Bert,” in *2022 7th International Conference on Image, Vision and Computing (ICIVC)*, Xi’an, China: IEEE, Jul. 2022, pp. 934–938. doi: 10.1109/ICIVC55077.2022.9886134.
- [18] S. Bano and S. Khalid, “BERT-based Extractive Text Summarization of Scholarly Articles: A Novel Architecture,” in *2022 International Conference on Artificial Intelligence of Things (ICAIoT)*, Istanbul, Turkey: IEEE, Dec. 2022, pp. 1–5. doi: 10.1109/ICAIoT57170.2022.10121826.
- [19] H. M. Liu and J. G. Lu, “Brief Survey of K-Means Clustering Algorithms,” *Appl. Mech. Mater.*, vol. 740, pp. 624–628, Mar. 2015, doi: 10.4028/www.scientific.net/AMM.740.624.
- [20] A. Venkataramana, K. Srividya, and R. Cristin, “Abstractive Text Summarization Using BART,” in *2022 IEEE 2nd Mysore Sub Section International Conference (MysuruCon)*, Mysuru, India: IEEE, Oct. 2022, pp. 1–6. doi: 10.1109/MysuruCon55714.2022.9972639.
- [21] M. Lewis *et al.*, “BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension,” in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, Online: Association for Computational Linguistics, 2020, pp. 7871–7880. doi: 10.18653/v1/2020.acl-main.703.
- [22] V. Rockova and E. Saha, “On Theory for BART,” 2018, *arXiv*. doi: 10.48550/ARXIV.1810.00787.
- [23] M. L. Adjie, W. G. Pribadi, L. A. Wulandhari, and G. Z. Nabiilah, “Improving Abstractive News Summarization Using Keyword Extraction for Human-Like Summaries,” in *2025 IEEE International Conference on Industry 4.0, Artificial*

- Intelligence, and Communications Technology (IAICT)*, Bali, Indonesia: IEEE, Jul. 2025, pp. 591–598. doi: 10.1109/IAICT65714.2025.11101587.
- [24] Z. Luo, “Knowledge-guided Aspect-based Summarization,” in *2023 International Conference on Communications, Computing and Artificial Intelligence (CCCAI)*, Shanghai, China: IEEE, Jun. 2023, pp. 17–22. doi: 10.1109/CCCAI59026.2023.00012.
- [25] A. Mastropaolo, M. Ciniselli, M. Di Penta, and G. Bavota, “Evaluating Code Summarization Techniques: A New Metric and an Empirical Characterization,” 2023, *arXiv*. doi: 10.48550/ARXIV.2312.15475.
- [26] G. W. Wicaksono, S. F. Al Asqalani, Y. Azhar, N. P. Hidayah, and A. Andreawana, “Automatic Summarization of Court Decision Documents over Narcotic Cases Using BERT,” *JOIV Int. J. Inform. Vis.*, vol. 7, no. 2, p. 416, May 2023, doi: 10.30630/joiv.7.2.1811.
- [27] M. S. Islam and L. Zhang, “A Review on BERT: Language Understanding for Different Types of NLP Task,” Jan. 26, 2024, *Computer Science and Mathematics*. doi: 10.20944/preprints202401.1857.v1.
- [28] M. Salıcı and Ü. E. Ölçer, “Impact of Transformer-Based Models in NLP: An In-Depth Study on BERT and GPT,” in *2024 8th International Artificial Intelligence and Data Processing Symposium (IDAP)*, Malatya, Turkiye: IEEE, Sep. 2024, pp. 1–6. doi: 10.1109/IDAP64064.2024.10710796.
- [29] P. Maharani Isnainiyah, C. S. K. Aditya, and D. R. Chandranegara, “Penerapan Model Pre-Trained BERT dalam Mendeteksi Teks Buatan ChatGPT,” *J. Repos.*, vol. 7, no. 3, pp. 393–400, Sep. 2025, doi: 10.22219/repositor.v7i3.37082.
- [30] J. Wang *et al.*, “Utilizing BERT for Information Retrieval: Survey, Applications, Resources, and Challenges,” *ACM Comput. Surv.*, vol. 56, no. 7, pp. 1–33, Jul. 2024, doi: 10.1145/3648471.
- [31] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” in *Proceedings of the*

- 2019 *Conference of the North*, Minneapolis, Minnesota: Association for Computational Linguistics, 2019, pp. 4171–4186. doi: 10.18653/v1/N19-1423.
- [32] K. P. Sinaga and M.-S. Yang, “Unsupervised K-Means Clustering Algorithm,” *IEEE Access*, vol. 8, pp. 80716–80727, 2020, doi: 10.1109/ACCESS.2020.2988796.
- [33] D. N. Alfiansyah, V. R. S. Nastiti, and N. Hayatin, “Penerapan Metode K-Means pada Data Penduduk Miskin Per Kecamatan Kabupaten Blitar,” *J. Repos.*, vol. 4, no. 1, Jan. 2024, doi: 10.22219/repositor.v4i1.31083.
- [34] Devaghi. P and S. Sudha, “Exploratory Data Analysis and Data Segmentation using K means Clustering,” in *2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, Greater Noida, India: IEEE, May 2023, pp. 145–147. doi: 10.1109/ICACITE57410.2023.10183143.
- [35] Mrs. J. Sirisha, V. Lakshmi Prathyusha, P. Naga Anupriya, M. Suma Sri, and P. Naga Hema, “Customer Segmentation using K-Means Clustering,” *Int. J. Adv. Res. Sci. Commun. Technol.*, pp. 170–175, Nov. 2022, doi: 10.48175/IJAR SCT-7618.
- [36] C.-Y. Lin, “ROUGE: A Package for Automatic Evaluation of Summaries,” *Assoc. Comput. Linguist.*, vol. Text Summarization Branches Out, pp. 74–81, Jul. 2004.
- [37] T. A. Mir, D. Banerjee, D. Upadhyay, and R. S. Rawat, “Binary Classification of Lung Diseases Using Convolutional Neural Networks and Support Vector Machines,” in *2024 4th Asian Conference on Innovation in Technology (ASIANCON)*, Pimari Chinchwad, India: IEEE, Aug. 2024, pp. 1–6. doi: 10.1109/ASIANCON62057.2024.10837890.
- [38] J. Chen, S. Wang, Z. Qi, Z. Zhang, C. Wang, and H. Zheng, “A Combined Encoder and Transformer Approach for Coherent and High-Quality Text Generation,” 2024, *arXiv*. doi: 10.48550/ARXIV.2411.12157.
- [39] B. M. Gurusamy, P. K. Rengarajan, and P. Srinivasan, “A hybrid approach for text summarization using semantic latent Dirichlet allocation and sentence concept

202110370311169
Fathul Agit Darmawan
Prodi Informatika

mapping with transformer,” *Int. J. Electr. Comput. Eng. IJECE*, vol. 13, no. 6, p. 6663, Dec. 2023, doi: 10.11591/ijece.v13i6.pp6663-6672.





FAKULTAS TEKNIK

INFORMATIKA

informatika.umm.ac.id | informatika@umm.ac.id

UNIVERSITAS MUHAMMADIYAH MALANG



FORM CEK PLAGIARISME LAPORAN TUGAS AKHIR

Nama Mahasiswa : FATHUL AGIT DARMAWAN
 NIM : 202110370311169
 Judul TA : Pendekatan Hibrida Peringkasan Transkripsi Video dengan Klasterisasi K-Means Berbasis BERT dan BART

Hasil Cek Plagiarisme dengan Turnitin

No.	Komponen Pengecekan	Nilai Maksimal Plagiarisme (%)	Hasil Cek Plagiarisme (%) *
1.	Bab 1 – Pendahuluan	10 %	6 %
2.	Bab 2 – Daftar Pustaka	25 %	12 %
3.	Bab 3 – Analisis dan Perancangan	25 %	0 %
4.	Bab 4 – Implementasi dan Pengujian	15 %	0 %
5.	Bab 5 – Kesimpulan dan Saran	5 %	0 %
6.	Makalah Tugas Akhir	20%	13 %

*) Hasil cek plagiarism diisi oleh pemeriksa (staf TU)

*) Maksimal 5 kali (4 Kali sebelum ujian, 1 kali sesudah ujian)

Mengetahui,
Pemeriksa (Staff TU)

(.....

.....)

