DATAMATIKA

Tersedia secara daring di:
https://conference.ipb.ac.id/datamatika

CLASSIFICATION OF INDONESIAN NEWS ARTICLES ON THE GENOCIDE OF PALESTINE USING LATENT DIRICHLET ALLOCATION AND DEEP LEARNING

KLASIFIKASI BERITA INDONESIA TERHADAP GENOSIDA PALESTINA MENGGUNAKAN *LATENT DIRICHLET ALLOCATION* DAN *DEEP LEARNING*

Salma Nadhira Danuningrat[1‡], Ahmad Ridha[2], and Sri Wahjuni[3]

[1,2,3]School of Data Science, Mathematics, and Informatics, IPB University, Indonesia
[1]Dicoding, Indonesia
‡corresponding author: salma.danuningrat@gmail.com

Abstract

Media classification in reporting on the Palestine-Israel conflict by Indonesian media plays a crucial role in shaping public perception and filtering out information bias. This study aims to automatically classify news related to the Palestinian genocide using a Natural Language Processing (NLP) approach. The dataset used consists of news headlines from six Indonesian online media outlets: Kompas, Tempo, Detik, Republika, CNN Indonesia, and CNBC Indonesia. Topic modeling using Latent Dirichlet Allocation (LDA) was applied as an initial step to heuristically define news categories. The results of this modeling were then used to label the data, which was then classified using deep learning models LSTM, Bi-LSTM, GRU, Bi-GRU, and IndoBERT. The results showed that from the six labels obtained from LDA, eight categories were developed for classification. The evaluation showed that IndoBERT performed best on manually labeled datasets with an average accuracy of 80,87% with stratified 5-fold cross-validation, while the model experienced overfitting on automatically labeled data obtained from LDA. Among the RNN models, LSTM performed best on the manually labeled dataset with an average accuracy of 67,62%, followed by GRU with an accuracy of 56,62%. These results also indicate that the unidirectional model better fits the relatively short characteristic headlines of news stories, compared to the bidirectional model, which tends to be too complex for this data context. This study emphasizes the importance of labeling quality, the potential of transformer-based methods

in understanding automatic media classification, and the relevance of NLP approaches in analyzing bias and issue representation in media reporting.

Keywords: Deep Learning, IndoBERT, Latent Dirichlet Allocation, Media Classification, Natural Language Processing, Palestinian Genocide.

Abstrak

Klasifikasi media dalam pemberitaan konflik Palestina-Israel oleh media Indonesia memiliki peran penting dalam membentuk persepsi publik dan menyaring bias informasi. Penelitian ini bertujuan untuk mengklasifikasikan berita terkait genosida Palestina secara otomatis dengan pendekatan *Natural Language Processing* (NLP). *Dataset* yang digunakan berupa judul berita dari enam media *online* Indonesia, yaitu Kompas, Tempo, Detik, Republika, CNN Indonesia, dan CNBC Indonesia. Pemodelan topik menggunakan *Latent Dirichlet Allocation* (LDA) diterapkan sebagai langkah awal untuk mendefinisikan kategori berita secara heuristik. Hasil pemodelan ini kemudian digunakan untuk melabeli data, yang selanjutnya diklasifikasikan menggunakan model *deep learning* LSTM, Bi-LSTM, GRU, Bi-GRU, dan IndoBERT. Hasil penelitian menunjukkan dari enam label hasil LDA, dikembangkan delapan kategori untuk klasifikasi. Evaluasi menunjukkan bahwa IndoBERT memiliki kinerja terbaik pada *dataset* berlabel manual dengan akurasi rata-rata sebesar 80,87% dengan *stratified* 5-*fold cross validation*, sementara model mengalami *overfitting* pada data berlabel otomatis hasil LDA. Di antara model RNN, LSTM menunjukkan performa terbaik pada *dataset* berlabel manual dengan akurasi rata-rata sebesar 67,62%, diikuti oleh GRU dengan akurasi sebesar 56,62%. Hasil ini juga menunjukkan bahwa model *unidirectional* lebih sesuai dengan karakteristik judul berita yang relatif pendek, dibandingkan dengan model *bidirectional* yang cenderung terlalu kompleks untuk konteks data ini. Penelitian ini menegaskan pentingnya kualitas pelabelan, potensi metode berbasis *transformer* dalam memahami klasifikasi media secara otomatis, serta relevansi pendekatan NLP dalam menganalisis bias dan representasi isu dalam pemberitaan media.

Kata Kunci: Deep Learning, Genosida Palestina, IndoBERT, Klasifikasi Media, Latent Dirichlet Allocation, Natural Language Processing.

1. Pendahuluan

Saat ini, internet penuh dengan berita *online* tentang berbagai topik yang mengandung informasi tidak relevan atau tidak bermanfaat. Oleh karena itu, penting untuk memproses konten berita *online* dan mengklasifikasikannya ke dalam berbagai kategori (Agarwal *et al.,* 2023). Klasifikasi topik berita adalah metode mengklasifikasikan artikel berita yang tersedia dalam data teks ke dalam beberapa kelas atau label yang telah ditentukan sebelumnya (Sunagar *et al.,* 2021).

Berbagai penelitian tentang klasifikasi media terkait konflik Israel-Palestina telah dilakukan. Ramadani *et al.* (2024) menunjukkan variasi pengkategoriasian dalam pemberitaan pengeboman Israel terhadap Palestina pada 1–3 November 2023: Detik, CNN Indonesia, dan Viva menyoroti korban dan kerusakan, sedangkan Kompas, Liputan6, dan CNBC menekankan narasi negatif terhadap Hamas sebagai pemicu konflik.

Klasifikasi teks secara manual adalah tugas yang menantang dan memakan waktu, terutama karena aturan klasifikasi yang harus dibuat secara manual dalam memproses banyaknya dokumen digital seperti berita *online* (Ahmed dan Ahmed, 2021). Pendekatan non-teknis, seperti literasi media atau analisis politik adalah langkah efektif untuk menghadapi tantangan bias berita, tetapi memerlukan banyak waktu dan usaha (Hamborg, 2023). Maka dari itu, proses klasifikasi secara otomatis, melalui teknik *Natural Language Processing* (NLP) mulai dikembangkan. Langkah pertamanya adalah penggunaan metode *unsupervised* seperti *topic modeling* untuk mengidentifikasi tema umum dalam kumpulan dokumen yang dapat merepresentasikan kategori. Salah satu pendekatan dalam pemodelan topik adalah Latent Dirichlet Allocation (LDA), sebuah model generatif probabilistik yang bekerja dengan cara menemukan struktur topik tersembunyi (*latent topics*) dari sekumpulan data teks (Chauhan dan Shah, 2021). Selanjutnya, topik tersebut dapat digunakan secara langsung atau dijadikan acuan untuk pelabelan manual. Topik yang sudah diberi label ini kemudian diterapkan metode *supervised* untuk mengidentifikasi kategori secara otomatis pada data yang belum berlabel (Nicholls, 2020) .

Secara umum, metode yang digunakan untuk mendeteksi kategori, seperti *topic modeling*, *clustering*, analisis sentimen, atau kombinasi dari metode tersebut mampu menghasilkan wawasan yang relevan dan tematik. Namun, hasilnya seringkali terlalu umum sehingga diragukan apakah metode tersebut benar-benar dapat disebut sebagai deteksi kategori (Jumle, 2024). Model berbasis *transformer* seperti IndoBERT terbukti menghasilkan tingkat akurasi yang lebih optimal dibandingkan dengan metode tradisional seperti *Support Vector Machine* (SVM) untuk klasifikasi teks (Prasetya *et al.,* 2023). Melalui metode *deep learning* dengan *neural network,* identifikasi kategori dari berbagai isu berita telah diterapkan. Menggunakan LDA bersamaan dengan BERT, Verbytska (2024) menemukan 18 topik dalam media timur dan 11 topik dalam media barat mengenai perang Rusia-Ukraina yang disimpulkan menjadi tujuh kategori utama. Liu *et al.* (2019) membandingkan *neural network* berbasis LSTM, GRU, dan BERT dalam mengidentifikasi *framing* berita Amerika Serikat terhadap topik kekerasan senjata. Model BERT berhasil mencapai akurasi 5-*fold cross validation* tertinggi sebesar 84,23% dalam mengidentifikasi sembilan jenis *frame* yang ditentukan berdasarkan literatur yang tersedia.

Klasifikasi pemberitaan media secara manual sering kali memakan waktu yang signifikan dan memerlukan pemahaman mendalam terkait topik yang dibahas. Penelitian bertujuan untuk mengembangkan *dataset* baru berisi judul-judul berita dari berbagai sumber berita Indonesia, mengevaluasi efektivitas metode *Latent Dirichlet Allocation* (LDA) untuk *topic modeling* pada *dataset* tersebut, dan membandingkan kinerja

beberapa model *deep learning* dalam mengklasifikasikan judul berita berdasarkan hasil *topic modeling* dengan LDA.

Penelitian ini diharapkan dapat berkontribusi pada pengembangan NLP berbahasa Indonesia yang masih terbatas dan bermanfaat dalam mendorong transparansi media dengan mengidentifikasi pola pemberitaan yang tidak berimbang sehingga dapat menjadi dasar untuk praktik jurnalistik yang lebih bertanggung jawab di Indonesia.

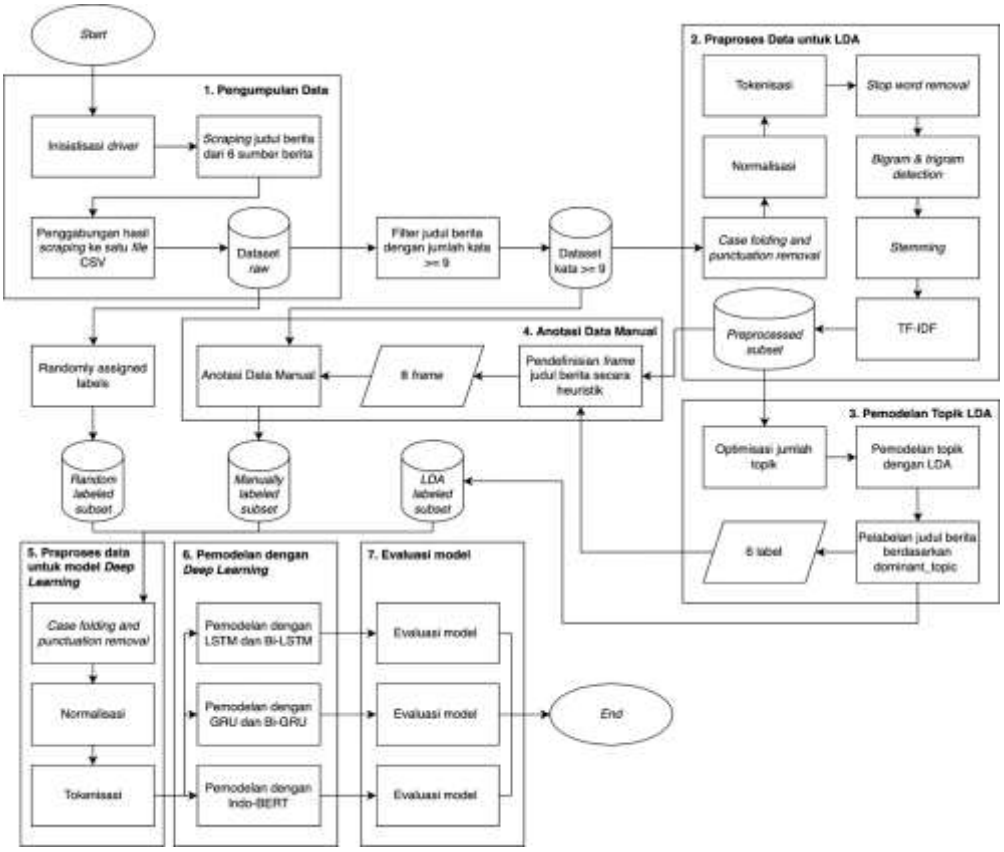
2. Metodologi

2.1 Data Penelitian

Judul berita merupakan elemen representatif utama yang mengandung fitur semantik paling menonjol, sehingga menjadi acuan krusial dalam mengidentifikasi dan mengklasifikasikan konsep suatu teks secara tepat (Sunagar *et al.,* 2021). Data yang digunakan dalam penelitian ini merupakan judul berita berbagai bentuk media (video, foto, infografik, dan artikel tertulis) yang diberi tagar #Palestina dan dipublikasi dalam rentang 471 hari, yaitu dari serangan kelompok militer Hamas pada 7 Oktober 2023 hingga kesepakatan tanggal permulaan gencatan senjata tanggal 19 Januari 2025. Judul berita diambil dari beberapa sumber media *online* yaitu Kompas, Tempo, Detik, Republika, CNN Indonesia, dan CNBC Indonesia. Pemilihan sumber berita tersebut bermaksud untuk membandingkan media dengan beragam nilai, segmen, dan lisensi.

2.2 Tahapan Penelitian

Penelitian ini dilaksanakan dalam tujuh tahapan yaitu pengumpulan data, praproses data untuk LDA, pemodelan topik dengan LDA, anotasi data manual, praproses data untuk model *deep learning*, pemodelan dengan *deep learning*, dan evaluasi model *deep learning*. Diagram alur dari tahapan penelitian terdapat pada Gambar 1.



Gambar 1: Tahapan penelitian.

2.3 Pengumpulan Data

Data didapatkan dengan proses *scraping* keenam situs web berita publik yang memiliki tagar #Palestina. *Scraping* data menggunakan Python dengan *library* BeautifulSoup4, Selenium WebDriver, serta SeleniumBase. Proses *scraping* data melalui empat tahapan, yakni inisialisasi *driver*, *scraping* data dengan mengumpulkan tautan yang berada dalam setiap halaman web lalu *scraping* isi dari setiap tautan yang dikumpulkan, serta penggabungan hasil *scraping* ke dalam satu *file* CSV. Hasil dari *scraping* data adalah satu *file* CSV yang berisi seluruh judul data yang dikumpulkan beserta atributnya yaitu: judul_berita, sumber_berita, url, dan tanggal_publikasi.

2.4 Praproses Data untuk LDA

Keberhasilan model yang adil dan transparan tidak dapat dicapai semata-mata melalui penyesuaian algoritma, melainkan harus tertanam kuat sejak awal siklus pemodelan AI, mulai dari akuisisi hingga prapemrosesan data (Karanam, 2024). Praproses data menggunakan bahasa pemrograman Python dengan memanfaatkan beberapa *library*. *Dataset* untuk LDA adalah *subset* dari *dataset* hasil *scraping* yaitu judul berita dengan jumlah kata sama dengan atau lebih dari sembilan kata. Tahapan praproses meliputi

penghilangan data duplikat, *case folding* dan penghapusan tanda baca dengan library Pandas, normalisasi, tokenisasi dengan library NLTK, pembersihan teks dari kata-kata yang termasuk dalam daftar *stopwords*, deteksi *bigram* dan *trigram,* menggunakan modul Phraser dalam library Gensim, *stemming* dengan library Sastrawi, serta *Term Frequency–Inverse Document Frequency* (TF-IDF) dengan model TfidfVectorizer dalam library Gensim.

2.5 Pemodelan Topik dengan LDA

Setelah data terkumpul dan di praproses, *topic modeling* diterapkan dengan metode LDA sebagai langkah *unsupervised* untuk menentukan topik-topik yang terkandung dalam data. Pemodelan topik dijalankan berulang kali dengan jumlah topik yang berbeda untuk mendapatkan model LDA optimal. Hasil akhir dari pemodelan LDA adalah sekumpulan topik yang masing-masing direpresentasikan oleh sekumpulan kata-kata dengan probabilitas kata dalam topik tersebut serta distribusi topik dalam setiap dokumen. Sejumlah topik tersebut kemudian dianalisis lebih lanjut dengan *library* pyLDAvis guna memvisualisasikan distribusi jarak antar topik.

Hasil LDA juga digunakan untuk memberikan label pada *subset* data berdasarkan *dominant topic*, yaitu topik dengan proporsi tertinggi dalam setiap dokumen. *Subset* ini dimanfaatkan untuk memvalidasi hasil pelabelan manual yang bersifat subjektif, dengan tujuan menilai sejauh mana model klasifikasi mampu menangkap konteks semantik yang didefinisikan oleh manusia.

2.6 Anotasi Data Manual

Berdasarkan kumpulan kata dalam setiap topik yang dihasilkan LDA, masing-masing topik diberikan definisi heuristik untuk dijadikan acuan dalam proses anotasi data manual. Data dianotasi secara mandiri dengan mencocokkan setiap judul berita dalam *dataset* dengan salah satu kategori.

2.7 Praproses Data untuk Model *Deep Learning*

Praproses data untuk model *deep learning* diterapkan kepada tiga variasi *dataset*, yaitu *dataset* dengan label manual, *dataset* dengan label otomatis LDA, dan *dataset* dengan label acak. Praproses meliputi tiga tahapan, yaitu *case folding* dan penghilangan tanda baca, normalisasi, dan tokenisasi.

2.8 Pemodelan dengan *Deep Learning*

Pemodelan beserta parameternya mengacu pada penelitian Liu *et al*. (2019) yang menggunakan tiga model *deep learning* untuk mendeteksi kategori dalam judul berita, yaitu *Long Short-Term Memory* (LSTM), *Gated Recurrent Unit* (GRU), dan BERT. Penelitian ini mengalami modifikasi dalam penggunaan IndoBERT, yaitu model BERT yang dilatih

menggunakan korpus Bahasa Indonesia. Setiap model dilatih menggunakan ketiga *dataset* yang akan diuji dengan pembagian data *stratified* 5-*fold cross validation*.

2.9 Evaluasi Model *Deep Learning*

Seluruh model yang dilatih dengan ketiga *dataset* dievaluasi berdasarkan akurasi dan F1-*score* untuk setiap labelnya. Akurasi dihitung sebagai proporsi prediksi yang benar terhadap jumlah total data uji. F1-*score* merupakan rata-rata harmonik dari *precision* dan *recall* sehingga memberikan gambaran keseimbangan antara kedua metrik tersebut.

3. Hasil dan Pembahasan

3.1 Pengumpulan Data

Data yang terkumpul dari keenam sumber berita sejumlah 23.384 judul berita. Atribut data yang dikumpulkan berupa judul_berita, sumber_berita, url, tanggal_publikasi.

Hasil *scraping* dari keenam sumber berita kemudian digabungkan ke dalam satu file Excel dan ditambahkan kolom jumlah_kata yang berisi jumlah kata dalam setiap judul_berita. Langkah ini guna membuat *subset* dari *dataset*, sehingga tahapan pemodelan selanjutnya tidak dipengaruhi oleh *noise* yang disebabkan oleh data yang terlalu pendek dan tidak menyumbangkan konteks yang bermakna.

3.2 Praproses Data untuk LDA

Pemodelan topik LDA dilakukan melalui dua tahap: pengembangan *pipeline* praproses dan penentuan jumlah topik. Untuk meminimalisir *noise*, penelitian menggunakan subset 17.390 judul berita (minimal sembilan kata) yang kemudian diproses melalui tahapan *case folding*, penghapusan tanda baca, normalisasi, tokenisasi, *stopwords removal*, deteksi bigram/trigram, *stemming*, hingga TF-IDF.

3.3 Pemodelan dengan LDA

Model LDA kemudian diuji dengan berbagai jumlah topik untuk menentukan jumlah topik yang paling sesuai berdasarkan tingginya nilai *coherence* dan *intertopic distance map* yang merepresentasikan jarak antar topik dalam ruang vektor. Berdasarkan dua pertimbangan tersebut, jumlah topik yang ditetapkan untuk langkah selanjutnya adalah enam topik.

Enam topik yang dihasilkan oleh model LDA, beserta kata-kata dengan bobot terbesar terhadap topiknya (*topic importance* tertinggi) dalam masing-masing topik, dijadikan acuan dalam mendefinisikan label kategori secara heuristik, yang terlihat dalam Gambar 2. Kumpulan kategori tersebut kemudian digunakan dalam proses pelabelan manual yang

selanjutnya dimanfaatkan untuk membandingkan performa berbagai model *deep learning* dalam mengklasifikasi.

Gambar 2: *Word cloud* untuk 6 topik berdasarkan kata dengan *importance* tertinggi.

3.4 Anotasi Data Manual

Melanjutkan hasil dari pemodelan topik, *subset* data dengan jumlah kata dalam judul berita lebih atau sama dengan sembilan kata diberi label berdasarkan kategori yang didefinisikan oleh peneliti secara heuristik dari hasil pemodelan topik LDA. Dikembangkan delapan kategori dari enam topik hasil LDA, yaitu kategori “Kemanusiaan”, “Perang dan Peran Amerika Serikat”, “Sikap Indonesia”, “Korban Tewas dan Eskalasi Konflik”, “Hamas dan Kelompok Bersenjata”, “Keterlibatan Negara Arab dan Politik Regional”, “Respon Dunia, Budaya Populer, Kegiatan Religi”, dan “Diplomasi Internasional”.

Dalam proses pelabelan manual, judul-judul berita yang tidak relevan dengan fokus isu dihapus dari *dataset* untuk menjaga konsistensi dan relevansi konteks dalam proses pelabelan. Hasil akhir dari anotasi data manual berupa 9.346 judul berita yang diberi label 0 hingga 7.

3.5 Praproses Data untuk Model *Deep Learning*

Praproses untuk klasifikasi dengan model *deep learning* meliputi case folding, normalisasi, dan tokenisasi yang difokuskan untuk menjaga struktur serta konteks kalimat. Hal ini berbeda dengan LDA yang menggunakan pendekatan bag-of-words untuk menyederhanakan bentuk kata demi meningkatkan koherensi topik tanpa memperhatikan urutan kata.

3.6 Pemodelan dengan *Deep Learning*

Setelah *dataset* yang telah diberi label kategori dari penentuan kategori secara heuristik dipraproses, klasifikasi kategori diuji dengan membandingkan lima model *deep learning*, yaitu LSTM, Bi-LSTM, GRU, Bi-GRU, dan IndoBERT.

Model LSTM (128 neuron) dan GRU (64 neuron) menggunakan arsitektur *Sequential* dan *Embedding* dengan *dropout* 0,2 serta regularisasi

L1 untuk mencegah *overfitting*. Model GRU ditambahkan SpatialDropout1D dan lapisan Dense (ReLU) sebelum aktivasi *softmax*. Versi *bidirectional* pada kedua model diterapkan untuk menangkap konteks dua arah.

Sementara itu, IndoBERT menggunakan model *pre-trained* "indolem/indobert-base-uncased" dengan lapisan *dropout* (0,3) dan Linear pada token [CLS]. Seluruh model dilatih menggunakan *stratified* 5-*fold cross validation*, di mana IndoBERT secara spesifik menggunakan optimizer AdamW dengan *learning rate* 2e-5 dan *batch size* 16.

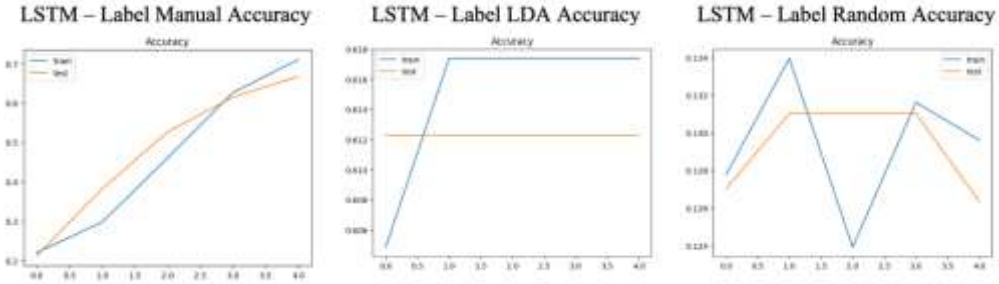
3.7 Evaluasi Model *Deep Learning*

Data Di antara model RNN yang diuji pada *dataset* label manual, LSTM menunjukkan performa terbaik dengan akurasi rata-rata sebesar 0,67. Model *unidirectional* LSTM dan GRU menunjukkan kinerja lebih unggul dibandingkan versi *bidirectional*. Hal ini diduga karena karakteristik judul berita yang cenderung singkat dan memuat informasi utama di awal kalimat, sehingga penambahan konteks dari arah sebaliknya justru meningkatkan kompleksitas yang tidak diperlukan. Hasil *accuracy* dan *loss* untuk seluruh model dapat dilihat pada Tabel 1.

Tabel 1: Hasil *accuracy* dan *loss* dengan *stratified* 5-*fold cross validation.*

Model	Accuracy	Loss
LSTM	0,6762 $\pm$ 0,0119	1,0660 $\pm$ 0,0234
Bi-LSTM	0,6611 $\pm$ 0,0185	1,1079 $\pm$ 0,0487
GRU	0,5662 $\pm$ 0,0210	1,2944 $\pm$ 0,0254
Bi-GRU	0,5233 $\pm$ 0,0234	1,3913 $\pm$ 0,0595
IndoBERT	0,8087 $\pm$ 0,0094	1,5271 $\pm$ 0,0157

IndoBERT mencapai akurasi tertinggi (0,81) pada *dataset* berlabel manual berkat mekanisme *self-attention* dan penggunaan *pre-trained embedding* yang mampu menangkap konteks semantik secara mendalam. F1 rata-rata makro tertimbang, yaitu jumlah dari perkalian F1-*score* untuk setiap label dengan jumlah judul berita pada label tersebut lalu dibagi banyaknya data, adalah sebesar 0,812. Sebagai validasi terhadap subjektivitas pelabelan, pengujian pada *dataset* berlabel LDA dan acak membuktikan bahwa label manual memiliki kualitas semantik yang lebih baik untuk dipelajari model tanpa risiko *overfitting*. Sebaliknya, meskipun model pada *dataset* LDA menunjukkan akurasi tertentu, grafik akurasi (Gambar 3) mengindikasikan adanya gejala *overfitting* yang signifikan dibandingkan dengan *dataset* manual.



Gambar 3: Hasil akurasi ketiga *dataset* pada LSTM.

Dugaan bahwa model mengalami *overfitting* semakin diperkuat dengan penurunan akurasi signifikan pada model IndoBERT saat diuji pada *dataset* LDA dibandingkan *dataset* manual, yakni dengan selisih akurasi sebesar 0,1436. Seluruh hasil akurasi dari kelima model pada ketiga jenis *dataset* dirangkum dalam Tabel 2.

Tabel 2: Hasil *accuracy stratified 5-fold cross validation* ketiga *dataset.*

Dataset	LSTM	Bi-LSTM	GRU	Bi-GRU	IndoBERT
Manual	0,6762	0,6611	0,5662	0,5233	0,8087
LDA	0,6175	0,6737	0,6514	0,6554	0,6651
Random	0,1296	0,1282	0,1299	0,1296	0,1264

4. Simpulan dan Saran

Penelitian ini mengumpulkan 23.384 judul berita dan menetapkan enam topik optimal melalui LDA dengan nilai coherence 0,508. IndoBERT meraih performa tertinggi pada *dataset* manual dengan nilai F1 rata-rata makro tertimbang sebesar 0,812, membuktikan kemampuannya menangkap konteks semantik meskipun pelabelan bersifat subjektif. Di antara model RNN, LSTM merupakan yang terbaik, diikuti Bi-LSTM dan GRU. Sebaliknya, klasifikasi dengan label otomatis LDA memicu *overfitting* dan akurasi rendah (0,665), sehingga LDA dinilai kurang efektif dalam mengidentifikasi topik pada *dataset* ini.

Saran untuk penelitian selanjutnya mencakup validasi anotasi melalui *inter-annotator agreement* dengan pakar, penerapan *multi-label classification*, serta eksplorasi *hyperparameter* lebih lanjut pada model LDA.

Daftar Pustaka

Ahmed A., Ahmed A. (2021). Online news classification using machine learning techniques. *IIUM Engineering Journal*. 22(2):209–221. doi: 10.31436/iiumej.v22i2.1662.

Agarwal J., Christa S., H. P., & Kumar M., S. G. (2023). Machine Learning Application for News Text Classification. Di dalam: 2023 13th International Conference on Cloud Computing, Data Science & Engineering (Confluence). hlm 463–466. doi: 10.1109/confluence56041.2023.10048856.

Chauhan P., Shah N. (2021). Topic Modeling Using Latent Dirichlet Allocation: A Survey. Di dalam: 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS). hlm 381–386. doi: 10.1109/ICCCIS51004.2021.9487739.

Hamborg F. (2023). Revealing Media Bias in News Articles. *Springer Nature*. https://link.springer.com/book/10.1007/978-3-031-17693-7.

Jumle V., Makhortykh M., Sydorova M., & Vziatysheva V. (2024). Finding frames with BERT: A transformer-based approach to generic news frame detection. ArXiv (Cornell University). doi: 10.48550/arxiv.2409.00272.

Karanam U.K. (2024). Data Biases and Data Quality in AI. *Iowa State University Digital Repository*.

Liu S., Guo L., Mays K., Betke M., & Wijaya D.T. (2019). Detecting Frames in News Headlines and Its Application to Analyzing News Framing Trends Surrounding U.S. Gun Violence. Di dalam: Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL). pp. 504–514. doi: 10.18653/v1/k19-1047.

Nicholls T., Culpepper P.D. (2020). Computational Identification of Media Frames: Strengths, Weaknesses, and Opportunities. Political Communication. 1–23. doi: 10.1080/10584609.2020.1812777.

Prasetya H.A., Lhaksamana K.M., & Al Faraby S. (2023). Perbandingan Metode SVM dan IndoBERT untuk Klasifikasi Teks pada Dokumen Legal. Jurnal RESTI (Rekayasa Sistem dan Teknologi). 7(4):855–862. doi: 10.29207/resti.v7i4.5021.

Ramadani M.S., Kurniawan K., & Fuadin A. (2024). Menguak Bias Media dalam Pemberitaan Konflik Israel-Palestina: Sebuah Analisis Konten Kritis. *Onoma*. 10(1):887–905. doi:10.30605/onoma.v10i1.3392.

Siswanti N. (2019). Analisis Framing Media: Studi Komparatif Media Online “CNN” dan “Kompas” Terkait Fenomena Kemanusiaan di Al-Aqsa Periode 20-23 Juli 2017. Jurnal Riset Komunikasi. 2(2):110–125.

Sunagar P., Kanavalli A., Nayak S., Mahan S., Prasad S., & Prasad S. (2021). News Topic Classification Using Machine Learning Techniques. hlm 461–474. https://doi.org/10.1007/978-981-33-4909-4_35.

Verbytska A. (2024). Topic Modelling as a Method for Framing Analysis of News Coverage of the Russia-Ukraine War in 2022–2023. *Language & Communication*. 99:174–193. doi: 10.1016/j.langcom.2024.10.004.