

Pemodelan Isu Perubahan Iklim di Indonesia Berdasarkan *Headline* Media

Anang Kunaefi¹, Aris Fanani²

¹ Program Studi Sistem Informasi, Fakultas Sains dan Teknologi, UIN Sunan Ampel Surabaya, Surabaya, Jawa Timur 60294, Indonesia

² Program Studi Matematika, Fakultas Sains dan Teknologi, UIN Sunan Ampel Surabaya, Surabaya, Jawa Timur 60294, Indonesia

[Diajukan: 1 November 2023, Direvisi: 19 Januari 2024, Diterima: 25 April 2024]

Penulis Korespondensi: Anang Kunaefi (email: akunaefi@uinsa.ac.id)

INTISARI — Perubahan iklim telah menjadi isu global yang memengaruhi semua negara dalam beberapa dekade terakhir. Fenomena ini menimbulkan kekhawatiran bagi Indonesia yang notabene adalah salah satu episentrum perubahan iklim. Berbagai penelitian menunjukkan bahwa perubahan iklim dapat merugikan masyarakat, seperti produksi pertanian yang tidak stabil, penurunan kesehatan masyarakat, dan pemanasan global. Penelitian ini mencoba memodelkan dan menganalisis topik perubahan iklim yang dibahas di media. Menemukan topik-topik tersembunyi dari teks dapat memberikan petunjuk dan informasi mengenai perbincangan publik seputar perubahan iklim, seperti pemikiran, persepsi, dan kesiapan masyarakat dalam memitigasi kemungkinan dampak buruk perubahan iklim. Untuk mengidentifikasi subjek tersembunyi dari korpus, penelitian ini memodelkan isu perubahan iklim di Indonesia menggunakan algoritma *latent Dirichlet allocation* (LDA) untuk menganalisis teks dari berita utama media di Indonesia. Sebanyak 7.000 data *headline* dari lima media daring dikumpulkan sepanjang tahun 2017 hingga 2021 dengan menggunakan teknik *web scraping*. Dengan ketentuan nilai koherensi tertinggi sebesar 0,560, pendekatan yang diusulkan menghasilkan delapan topik terkait perubahan iklim. Topik-topik tersebut adalah energi terbarukan, emisi karbon, pengelolaan lingkungan hidup, ekonomi pembangunan, kerja sama internasional, kebijakan/regulasi, rehabilitasi, dan bencana alam. Berdasarkan hasil yang diperoleh, model tersebut cukup mampu menggambarkan tema pembahasan di masyarakat dan menangkap pemikiran masyarakat serta kesiapan pemerintah dalam bentuk kebijakan dan regulasi untuk menghadapi fenomena perubahan iklim.

KATA KUNCI — Perubahan Iklim, LDA, Analisis Teks, Penambangan Teks, Pemodelan Topik.

I. LATAR BELAKANG

Perubahan pola iklim yang disebabkan oleh emisi gas rumah kaca dan aktivitas manusia lainnya yang mengabaikan tatanan alam disebut sebagai perubahan iklim. Perubahan tersebut menimbulkan berbagai fenomena alam yang merugikan manusia, antara lain bencana alam dan pemanasan global. Secara umum, telah terjadi peningkatan emisi gas dan suhu sebesar 1 °C–1,5 °C di seluruh dunia. Pada tahun 2018, terdapat 315 kasus bencana alam akibat perubahan iklim yang menimpa lebih dari 68,5 juta orang, dengan kerugian materi mencapai US\$31,7 miliar [1]. Selain itu, ketahanan pangan, air, kesehatan, dan ekosistem rentan terhadap perubahan iklim.

Banyak negara telah mengakui bahwa perubahan iklim mempunyai dampak negatif. World Meteorological Organization (WMO) dan Perserikatan Bangsa-Bangsa mendirikan Intergovernmental Panel on Climate Change (IPCC) pada tahun 1988 (IPCC 2013) untuk mendorong semua negara agar memiliki pengetahuan yang sama ketika menetapkan kebijakan mengenai perubahan iklim. Pembentukan United Nations Framework Convention on Climate Change (UNFCCC) pada tahun 1992 merupakan langkah yang paling penting. Konvensi ini telah disetujui oleh 169 negara di seluruh dunia. Conference of the Parties (COP) diadakan setiap tahun oleh perwakilan negara-negara anggota UNFCCC untuk membahas kemajuan dalam perjuangan melawan perubahan iklim. Berlin, Jerman, menjadi tuan rumah COP UNFCCC pertama pada tahun 1995.

Indonesia, sebagai negara penghasil emisi gas rumah kaca terbesar ketiga di dunia, dipandang sebagai negara yang rentan terhadap perubahan iklim [2]. Berbagai penelitian menunjukkan bahwa perubahan iklim memberikan dampak

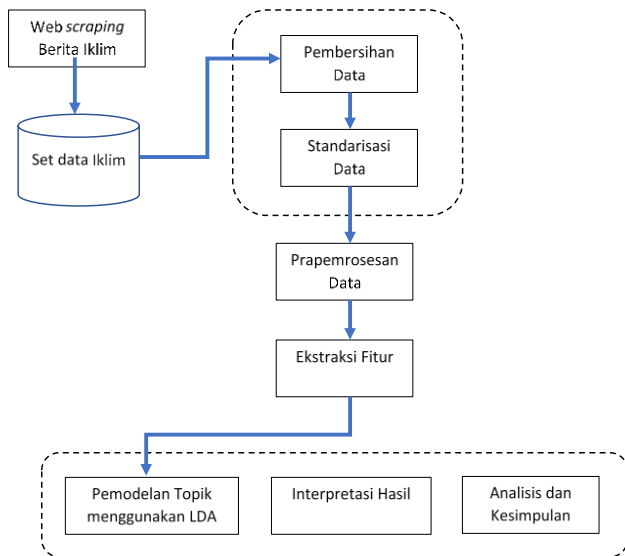
negatif, seperti penurunan produksi pertanian [3], [4]; kerentanan ketahanan pangan [5]; penggundulan hutan; dan ancaman terhadap keanekaragaman hayati lingkungan [6].

Oleh karena itu, diperlukan upaya yang sistematis dari seluruh pemangku kepentingan, terutama pemerintah, dalam menghadapi tantangan perubahan iklim. Tingkat perhatian dan kesadaran para pemangku kepentingan di Indonesia terhadap fenomena perubahan iklim dapat dilihat dan dievaluasi dari pemberitaan di media yang setiap hari menjadi perbincangan di masyarakat.

Penelitian ini mencoba memodelkan dan menganalisis topik-topik perubahan iklim yang dibahas di media untuk menggambarkan pemikiran dan persepsi publik dalam menghadapi ancaman-ancaman yang telah disebutkan sebelumnya. Selain itu, penelitian ini juga berusaha untuk mengetahui tantangan perubahan iklim dan melihat ada tidaknya upaya antisipasi wacana publik, khususnya dari para pembuat kebijakan di Indonesia.

Penelitian sebelumnya menggunakan pendekatan penambangan teks untuk memodelkan tema perubahan iklim dari teks media. Sebagai contoh, sebuah studi memodelkan perubahan iklim berdasarkan liputan media berita di India [7], sementara studi lain menganalisis kebijakan mengenai perubahan iklim menggunakan algoritma pemrosesan bahasa alami [8]. Namun dalam studi ini, pendekatan pemodelan masalah perubahan iklim telah diusulkan bersama dengan *pipeline* pembelajaran mesin dan hasilnya didiskusikan.

Dalam penelitian ini, set data dibangun dengan cara *scraping* teks *headline* media dari beberapa situs web media. Set data kemudian dibersihkan, melalui prapemrosesan, dan disimpan untuk algoritma *clustering*. Selanjutnya, algoritma *latent Dirichlet allocation* (LDA) digunakan untuk



Gambar 1. Metodologi penelitian.

menganalisis teks, sehingga diperoleh pengetahuan tersembunyi dalam korpus. Alasan penggunaan LDA dibandingkan algoritma lain adalah kecocokan LDA untuk kumpulan teks pendek, seperti judul *headline* dan ulasan pengguna.

Menganalisis konten media dapat memberikan petunjuk dan informasi mengenai topik perubahan iklim yang sedang dibicarakan di masyarakat. Lebih lanjut, isu yang muncul berdasarkan model penambangan teks ini dapat menggambarkan kesiapan Indonesia dalam menghadapi tantangan perubahan iklim

Makalah ini memberikan dua kontribusi. Pertama, algoritma pemodelan topik diimplementasikan berdasarkan set data *headline* media Indonesia. Kedua, temuan diperiksa, dijelaskan, dan dibandingkan dengan penelitian sebelumnya.

Makalah ini tersusun atas: Bagian I menyajikan gambaran umum penelitian ini. Pada Bagian II, metodologi dijelaskan secara menyeluruh, beserta set data dan desain eksperimennya. Hasil empiris percobaan disajikan pada Bagian III, kemudian pembahasan hasil dipaparkan pada Bagian IV. Bagian V mencakup rekomendasi untuk penyelidikan lebih lanjut.

II. METODOLOGY

Dalam penelitian ini, metode penelitian yang diusulkan terdiri atas empat langkah utama: a) pengumpulan data, b) prapemrosesan data, c) pemodelan topik, serta d) analisis dan diskusi topik, sebagaimana disajikan pada Gambar 1. Pada subbab berikutnya, setiap langkah dari metode ini dijelaskan.

A. PENGUMPULAN DATA

Langkah awal dalam alur kerja ini adalah mengumpulkan set data, yaitu berita utama artikel, dari beberapa sumber media nasional. Agar lebih sederhana, *headline* yang digunakan adalah *headline* dari surat kabar daring. Namun demikian, untuk menjamin validitas dan integritas data, beberapa kriteria diterapkan. Pertama, media daring yang dipilih harus memiliki versi cetak koran, atau media daring tersebut harus sudah dikenal publik dan memiliki reputasi yang diakui secara nasional. Kedua, media daring yang dipilih harus menyediakan akses ke artikel yang diterbitkan dari tahun 2017 hingga 2021, sesuai dengan kerangka waktu penelitian ini. Berdasarkan kriteria tersebut, media daring yang memenuhi syarat adalah Kompas, Detik, Merdeka, SindoNews, dan AntaraNews, seperti yang tercantum pada Tabel I.

TABEL I
SUMBER DATA MEDIA DARING

Nama Media	URL	Jumlah <i>Headline</i>
Kompas	www.kompas.com	290
Detik	www.detik.com	109
Merdeka	www.merdeka.com	50
SindoNews	www.sindonews.com	762
AntaraNews	www.antaranews.com	6.000

TABEL II
CONTOH HASIL PRAPEMROSESAN DATA

No.	Teks <i>Headline</i>	Hasil
1.	“Tanam Mangrove, Pemerintah Pulihkan Lingkungan Sekitar Pantai”	“tanam,” “mangrove,” “pemerintah,” “pulihkan,” “lingkungan,” “sekitar,” “pantai”
2.	“BMKG Tingkatkan Akurasi Informasi Cuaca Dengan Metode Baru”	“bmgk,” “tingkatkan,” “akurasi,” “informasi,” “cuaca,” “metode,” “baru”
3.	“Perubahan Iklim Perbesar Kerugian Ekonomi Para Petani”	“perubahan,” “iklim,” “perbesar,” “kerugian,” “ekonomi,” “petani”
4.	“Relawan Muda Untuk Bumi Yang Sehat Lakukan Penanaman Ribuan Pohon”	“relawan,” “muda,” “bumi,” “sehat,” “lakukan,” “penanaman,” “ribuan,” “pohon”

Selanjutnya digunakan teknik *web scraping*, yaitu teknik pencarian data pada situs web secara otomatis untuk mendapatkan *headline* pada situs web yang memenuhi kriteria tertentu. Frasa “perubahan iklim” digunakan sebagai kata kunci untuk memahami teks *headline*. Total data *headline* yang dikumpulkan adalah 7.000 *headline*.

B. PRAPEMROSESAN DATA

Tahap ini terdiri atas dua tahap pengolahan, yaitu pembersihan data dan prapemrosesan linguistik yang masing-masing akan dijelaskan sebagai berikut.

1) PEMBERSIHAN DATA

Data yang diperoleh dari prosedur *web scraping* seringkali mengandung banyak derau. Pasalnya, data teks yang dikumpulkan berasal dari banyak sumber yang formatnya beragam [9]. Oleh karena itu, pada tahap ini dilakukan pembersihan data agar data tersebut dapat digunakan untuk proses selanjutnya. Proses pembersihan data terdiri atas tiga langkah, yaitu menghilangkan teks kosong, menghilangkan karakter simbol, dan standarisasi data.

Dalam hal penghapusan teks kosong, data yang diambil dari web sering kali berisi teks kosong karena kesalahan aplikasi atau masalah teknis lainnya. Jika hal itu terjadi, rekaman (*record*) yang berisi teks kosong akan dihapus untuk memastikan integritas data.

Dalam pengumpulan data otomatis, biasanya teks berisi karakter simbolis yang tidak memberikan kontribusi terhadap makna teks. Beberapa contoh simbol adalah emotikon, tanda baca, atau simbol lainnya yang disebabkan oleh kesalahan penulisan atau perbedaan format teks. Simbol ini harus dihilangkan untuk menghindari gangguan pada proses penambangan teks. Sementara itu, sisa teks disimpan untuk langkah berikutnya.

Secara umum, proses standarisasi data bertujuan untuk membakukan format data hanya dengan menggunakan satu formulir untuk menghindari ambiguitas pada data. Proses ini penting karena data teks berasal dari berbagai sumber dengan bentuk yang berbeda-beda. Misalnya, data tanggal pada suatu situs web menggunakan format tanggal pendek (angka), sedangkan pada situs web lain ditulis dalam bentuk teks (huruf).

2) PRAPEMROSESAN LINGUISTIK

Tahap prapemrosesan linguistik bertujuan untuk mempersiapkan data sesuai dengan kebutuhan algoritma penambangan teks. Terdapat tiga tahap pemrosesan: *case folding*, tokenisasi, dan penghilangan *stop word*.

Tahap pertama adalah *case folding*. Tahap ini bertujuan untuk mengubah semua huruf besar dalam teks menjadi huruf kecil. Algoritma penambangan teks bergantung pada probabilitas istilah yang bergantung pada frekuensi kemunculan sebuah istilah atau kata. Sebagai contoh, "Aman" akan diubah menjadi "aman" untuk menunjukkan bahwa kedua istilah ini sama.

Langkah selanjutnya adalah penghilangan *stop word*. *Stop word*, seperti kata penghubung atau kata ganti objek, adalah kata-kata nonspesifik yang sering muncul dalam kalimat. Mengingat sifatnya yang tidak spesifik, *stop word* harus dihilangkan karena tidak mengandung makna yang signifikan dalam proses penambangan teks. Beberapa contoh kata yang dikategorikan sebagai *stop word* adalah "dan," "adalah," dan "meskipun." Digunakan *stop word* bahasa Indonesia dengan memanfaatkan pustaka Natural Language Toolkit (NLTK).

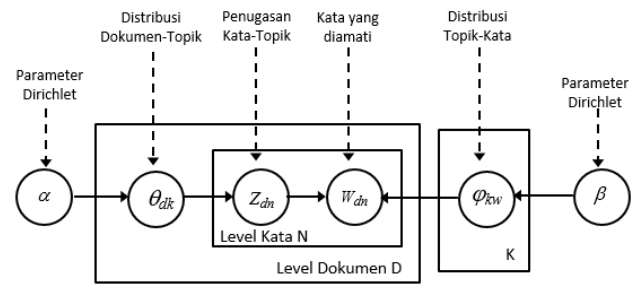
Langkah terakhir adalah tokenisasi. Proses tokenisasi adalah proses mengubah kalimat menjadi token (yaitu kata-kata). Langkah ini bertujuan untuk mempersiapkan data untuk algoritma pemodelan topik karena algoritma ini didasarkan pada daftar kata untuk setiap dokumen. Tabel II menyajikan beberapa contoh transformasi dari teks *headline* asli menjadi hasil dari langkah-langkah prapemrosesan data.

C. PEMODELAN TOPIK

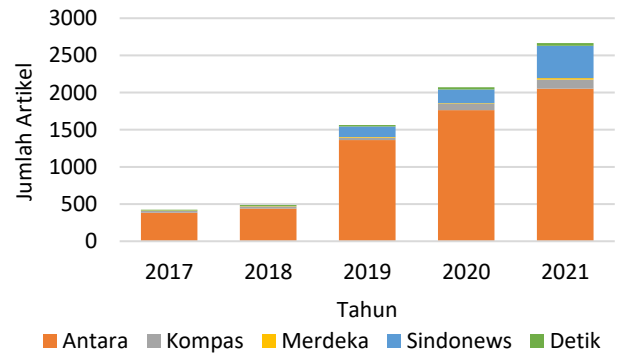
Pemodelan topik adalah teknik yang menonjol dalam bidang analisis teks dan pembelajaran mesin yang membantu dalam menemukan struktur tematik yang tersembunyi di dalam kumpulan dokumen. LDA adalah salah satu pendekatan untuk pemodelan topik yang paling sering digunakan [10].

Dasar dari LDA adalah bahwa setiap dokumen dalam sebuah korpus terdiri dari berbagai topik, dan setiap kata ditugaskan ke salah satu dari topik-topik ini. Hal ini menunjukkan bahwa terdapat sekumpulan topik yang mendasari seluruh korpus dan dokumen dihasilkan secara probabilistik berdasarkan topik-topik ini. Model ini menggunakan distribusi Dirichlet untuk memodelkan distribusi topik dalam dokumen dan distribusi kata dalam topik. Melalui proses generatif, LDA menetapkan topik ke dokumen dan kata ke topik secara berulang hingga menyatu menjadi sekumpulan topik yang koheren dan distribusi katanya yang terkait.

Model LDA telah banyak diaplikasikan dalam analisis teks, termasuk pengelompokan dokumen, dan pengelompokan ulasan singkat [11], [12]. Keserbagunaannya, kemampuan interpretasinya, dan kemampuannya untuk menemukan tema-tema tersembunyi membuatnya menjadi alat yang fundamental bagi para peneliti dan praktisi di bidang-bidang seperti pengambilan informasi, ilmu sosial, dan pemrosesan bahasa alami. Representasi intuitif topik sebagai distribusi probabilitas atas kata-kata telah membuat LDA menjadi alat yang sangat



Gambar 2. Model grafis LDA.



Gambar 3. Distribusi data berdasarkan frekuensi dan tahun.

diperlukan untuk memahami dan mengekstraksi informasi yang bermakna dari korpus teks yang besar.

LDA mengasumsikan bahwa terdapat K topik di seluruh korpus, yang mana K adalah parameter yang ditentukan pengguna. Setiap dokumen dianggap sebagai campuran dari topik-topik K ini, dan proporsi topik-topik ini dalam sebuah dokumen diatur oleh distribusi probabilitas, biasanya distribusi Dirichlet. Setiap topik dijelaskan dengan distribusi kata. Distribusi ini mencerminkan kemungkinan mengamati kata tertentu berdasarkan topiknya.

Berdasarkan model grafis LDA pada Gambar 2, saat membuat korpus, parameter Dirichlet sebelumnya (α dan β) diambil sampelnya satu kali pada tingkat korpus, dan parameter θ diambil sampelnya satu kali pada tingkat dokumen untuk setiap dokumen. Misalkan D adalah jumlah dokumen dalam korpus, W adalah jumlah kata unik dalam kosa kata, dan N adalah jumlah kata dalam suatu dokumen. Kemudian, θ adalah matriks $D \times K$ yang mewakili sebaran topik untuk setiap dokumen. θ_{dk} mewakili probabilitas topik k dalam dokumen d .

Parameter ϕ adalah matriks $K \times W$ yang mewakili distribusi kata untuk setiap topik, ϕ_{kw} mewakili probabilitas kata w dalam topik k , dan z adalah matriks $D \times N$ yang mewakili penetapan topik untuk setiap kata di setiap dokumen. Sementara itu, z_{dn} mewakili penetapan topik untuk kata ke- n dalam dokumen d , w adalah matriks $D \times N$ yang mewakili kata-kata dalam dokumen, dan w_{dn} mewakili kata ke- n dalam dokumen d .

III. HASIL

A. ANALISIS DATA EKSPLORASI

Analisis data eksplorasi dilakukan terhadap data untuk mendapatkan pemahaman awal set data yang lebih baik. Gambar 3 menunjukkan bahwa tahun 2021 memiliki set data terbesar dengan 2.666 *record*, sedangkan tahun 2017 memiliki set data terkecil dengan 425 *record*. Antara menempati peringkat teratas dalam penyebaran data, diikuti oleh Sindonews, Kompas, Merdeka, dan Detik, dengan persentase



Gambar 4. Word clouds berdasarkan tahun (a) 2018, (b) 2019, (c) 2020, dan (d) 2021.

yang sangat kecil dari total penyebaran data. Kebijakan yang mengatur jumlah data yang dapat diakses di masing-masing situs media berbeda-beda, hal ini menyebabkan variasi yang mencolok dalam distribusi data.

Kemudian, *word cloud* dihasilkan dari data tersebut, sehingga didapatkan kata yang paling sering muncul setiap tahunnya. Wordcloud adalah alat analisis yang berguna untuk melihat kata yang sering muncul di set data. Semakin besar tampilan istilah dalam *word cloud*, semakin besar pula frekuensi kata tersebut dalam set data.

Pada tahun 2018, isu yang paling sering muncul diwakili oleh kata “pemerintah,” “dunia,” “lingkungan,” dan “negara”. Gambar 4(a) menunjukkan frekuensi liputan media yang sering dilakukan pemerintah mengenai isu-isu lingkungan hidup.

Pada tahun 2019, seperti terlihat pada Gambar 4(b), isu yang ada di media tidak banyak berubah. Kata-kata yang sering muncul adalah “pemerintah,” “lingkungan,” “klhk” (Kementerian Lingkungan Hidup dan Kehutanan), dan “bmkg” (Badan Meteorologi, Klimatologi, dan Geofisika). Ilustrasi tersebut menunjukkan respons aktif pemerintah terhadap isu lingkungan hidup sehingga media mengutipnya. Kata kunci “bmkg” juga menunjukkan fenomena cuaca yang menonjol pada tahun tersebut.

Pada tahun 2020, terjadi pandemi COVID-19. Berdasarkan Gambar 4(c), pada tahun ini, istilah-istilah seperti “pemerintah,” “masyarakat,” “dunia,” “ekonomi,” dan “covid” sering disebutkan di media. Gambar tersebut menunjukkan bahwa pemberitaan media didominasi oleh pemberitaan mengenai pandemi COVID-19 yang mencapai puncaknya pada tahun 2020. Isu ini juga bersinggungan dengan masyarakat yang terkena dampak langsung PHK massal dan kesulitan keuangan lainnya.

Pada tahun 2021, sebagaimana ditunjukkan Gambar 4(d), pemberitaan media didominasi oleh kata “pemerintah,” “dunia,” “dampak,” dan “ekonomi.” Ilustrasi ini menunjukkan bahwa media kerap mengutip pernyataan pemerintah dalam mendorong pemulihan ekonomi sebagai dampak langsung dari pandemi.

B. IMPLEMENTASI PEMODELAN TOPIK

Pada tahap ini, pemodelan topik digunakan untuk menganalisis korpus. Gensim [13] dan pyLDAvis [14]

digunakan untuk mengimplementasikan algoritma pemodelan topik. Gensim adalah paket Python yang menyertakan algoritma LDA dan menyediakan implementasi *multi-core* yang lebih nyaman untuk meningkatkan kecepatan. Paket ini juga menyediakan fasilitas pemrosesan teks yang lebih bervariasi dibandingkan paket lainnya. PyLDAvis, di sisi lain, adalah paket yang dirancang untuk membantu pengguna menafsirkan topik yang diambil dari model topik LDA yang sesuai dan menyajikannya dalam visualisasi berbasis web yang interaktif.

Cara berinteraksi dengan data berdimensi tinggi dengan cara yang bermakna dan dapat dipahami manusia merupakan tantangan bagi banyak algoritma pembelajaran mesin, termasuk model LDA. Banyak peneliti menyatakan bahwa memilih hiperparameter secara heuristik dan kemudian menyempurnakannya menggunakan pengujian empiris merupakan strategi yang ideal untuk memecahkan masalah ini. Mengikuti strategi ini, tiga hiperparameter dalam algoritma LDA dipilih untuk pengaturan awal, dan nilainya disesuaikan berdasarkan hasil. Parameter yang disesuaikan adalah sebagai berikut.

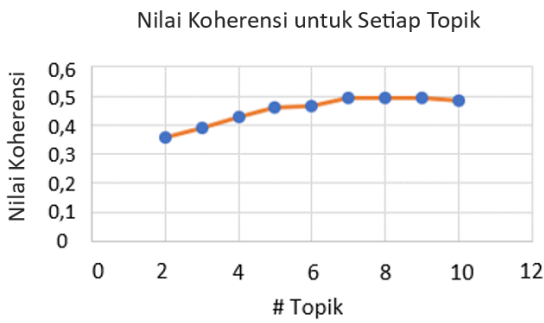
1) JUMLAH TOPIK

Jumlah topik sering kali menjadi salah satu hiperparameter terpenting dalam pemodelan topik. Namun, menentukan kelompok atau jumlah topik yang ideal sering kali merupakan suatu tantangan. Dalam percobaan ini, untuk menentukan jumlah topik yang optimal, algoritma diiterasi secara berurutan pada rentang nilai (yaitu, $n = 2, \dots, 10$) dari jumlah topik dan mengevaluasi hasilnya berdasarkan nilai koherensi.

Suatu pernyataan dikatakan koheren jika setiap kata dalam pernyataan tersebut saling mendukung. Dalam pemodelan topik, nilai koherensi dapat diperoleh dengan menghitung kemiripan antar kata yang sering muncul pada topik yang bersangkutan. Berdasarkan percobaan, jumlah kluster optimal adalah delapan (lihat Gambar 5), menghasilkan nilai koherensi rata-rata sebesar 0,494 dan nilai koherensi tertinggi sebesar 0,560 (lihat Tabel III).

2) HIPERPARAMETER α DAN β

Hiperparameter α menentukan kepadatan topik dokumen, sedangkan hiperparameter β menentukan kepadatan topik kata.



Gambar 5. Plot nilai koherensi untuk setiap topik.

TABEL III
NILAI KOHERENSI UNTUK SETIAP SKENARIO

No.	# Topik	Rerata Nilai Koherensi	Nilai Koherensi Tertinggi
1.	2	0,341	0,437
2.	3	0,398	0,450
3.	4	0,404	0,479
4.	5	0,438	0,524
5.	6	0,421	0,541
6.	7	0,447	0,559
7.	8	0,494	0,560
8.	9	0,486	0,556
9.	10	0,413	0,538

Dokumen dengan nilai α lebih tinggi berisi lebih banyak topik, sedangkan dokumen dengan nilai α lebih rendah memiliki tema lebih sedikit. Sebaliknya, topik terbentuk dari sejumlah besar kata dalam korpus ketika β tinggi dan dari beberapa kata ketika β rendah.

Dengan menggunakan nilai rentang pengaturan (yaitu 0,01...1) dari hiperparameter α dan β , algoritma diiterasi dengan peningkatan nilai sebesar 0,01 untuk setiap parameter guna memperoleh hasil optimal untuk setiap hiperparameter. Hasil nilai koherensi perbedaan nilai α dan β ditunjukkan pada Tabel IV. Akan tetapi, tidak semuanya dapat ditampilkan pada tabel karena banyaknya iterasi dan nilai hasil yang dihasilkan. Hanya sebagian kecil, yaitu hasil nilai tertinggi, yang ditampilkan pada Tabel IV karena pertimbangan kepraktisan. Terlihat pada tabel, nilai optimal hiperparameter α adalah 0,91, dan β adalah 0,01. Pengaturan ini digunakan untuk membangun model akhir, dan hasilnya akan dibahas.

IV. PEMBAHASAN

Berdasarkan nilai jumlah topik, yaitu delapan, hasil topik dari model tersebut dicantumkan dalam subbab berikut. Kemudian, interpretasi topik berdasarkan kata-kata yang muncul di setiap kluster akan dibahas.

A. INTERPRETASI TOPIK

1) TOPIK 1 (ENERGI TERBARUKAN)

Topik 1 didominasi oleh kata-kata seperti “energi,” “investasi” “esdm” (yaitu kementerian yang membidangi sumber daya mineral dan energi), dan “terbarukan.” Kategori ini dapat diartikan dengan topik energi terbarukan. Distribusi kata berdasarkan topik dapat dilihat pada Gambar 6(a). Topik energi erat kaitannya dengan isu lingkungan hidup, mengingat sebagian besar energi yang digunakan saat ini berbasis fosil. Kelemahan energi berbasis fosil terletak pada keterbatasan ketersediaannya yang terus menipis. Apalagi produksi energi berbasis fosil membutuhkan waktu puluhan hingga ratusan tahun.

TABEL IV
NILAI KOHERENSI α DAN β YANG BERBEDA

No.	Topik	Alpha (α)	Beta (β)	Nilai Koherensi
1.	8	0,01	0,01	0,542
2.	8	0,01	0,3	0,488
3.	8	0,01	0,04	0,449
4.	8	0,01	0,91	0,413
5.	8	0,31	0,01	0,549
6.	8	0,31	0,31	0,490
7.	8	0,31	0,04	0,457
8.	8	0,31	0,91	0,476
9.	8	0,04	0,01	0,549
10.	8	0,04	0,04	0,449
11.	8	0,04	0,31	0,517
12.	8	0,04	0,91	0,452
13.	8	0,91	0,01	0,560
14.	8	0,91	0,31	0,487
15.	8	0,91	0,04	0,436
16.	8	0,91	0,91	0,430

Topik energi terbarukan juga berkaitan dengan perubahan iklim. Meningkatnya suhu bumi, yang merupakan fenomena perubahan iklim global, membuat para peneliti untuk mempertimbangkan penggunaan energi berbasis surya (misalnya panel surya). Hasil ini sesuai dengan [15] yang menyatakan bahwa pengembangan energi terbarukan, seperti energi berbasis surya dapat memberikan dampak positif terhadap perubahan iklim.

2) TOPIK 2 (EMISI KARBON)

Pada topik ini muncul kata-kata seperti “emisi,” “karbon,” “industri,” “kurang,” “konservasi,” “teknologi,” “hijau,” dan beberapa kata lainnya. Dapat diartikan bahwa pada kategori ini, model tersebut mengemukakan topik pengurangan emisi atau pengurangan gas karbon, khususnya yang disebabkan oleh industri.

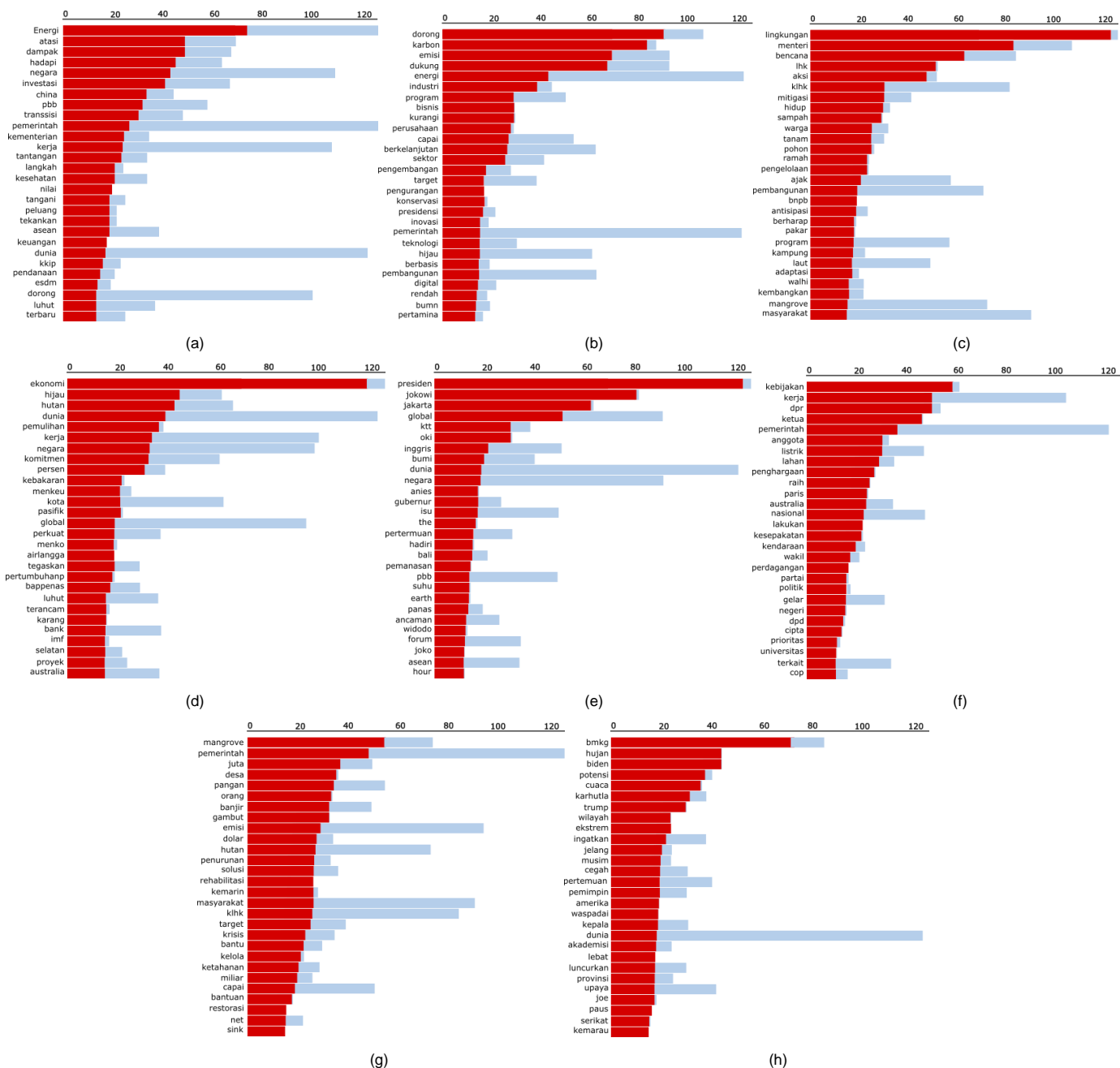
Tidak dapat dipungkiri bahwa peningkatan emisi gas karbon yang berasal dari industri atau eksploitasi skala besar turut menjadi penyebab terjadinya fenomena perubahan iklim. Oleh karena itu, perlu ada aksi masif untuk mengurangi jumlah emisi karbon. Distribusi kata pada kelompok ini dapat dilihat pada Gambar 6(b).

Topik emisi karbon dapat ditangkap dengan algoritma LDA. Algoritma ini juga mengelompokkan emisi karbon dengan kata-kata seperti “pengurangan” dan “konservasi” dalam kategori yang sama. Hasil ini menunjukkan bahwa telah ada upaya penurunan emisi dari berbagai pihak, salah satunya melalui konservasi dan penggunaan teknologi hijau yang semakin meluas. Hasil ini selaras dengan saran dalam [16], bahwa pengurangan emisi secara cepat harus dilakukan untuk memulihkan ekosistem.

3) TOPIK 3 (PENGELOLAAN LINGKUNGAN)

Dalam kategori ini, kata-kata yang muncul dalam kluster adalah “lingkungan,” “tanam,” “pohon,” “sampah,” “walhi” (yaitu lembaga swadaya masyarakat Indonesia yang berfokus pada isu lingkungan hidup), “mangrove,” dan lain-lain. Dapat diartikan bahwa tema pada topik 3 adalah pengelolaan lingkungan hidup. Distribusi kata pada topik 3 dapat dilihat pada Gambar 6(c).

Kata “lingkungan” meraih skor tertinggi pada kategori ini. Algoritma tersebut dapat mengklasifikasikan aktivitas



Gambar 6. Visualisasi LDA untuk setiap topik, (a) topik 1, (b) topik 2, (c) topik 3, (d) topik 4, (e) topik 5, (f) topik 6, (g) topik 7, dan (h) topik 8. Sumbu Y berisi 30 istilah paling relevan untuk setiap topik. Sumbu X menggambarkan jumlah setiap suku.

pengelolaan lingkungan yang melibatkan manusia dan komunitas ke dalam satu kategori topik. Pengelolaan lingkungan hidup erat kaitannya dengan perubahan iklim, yang mana kegiatan seperti penanaman pohon dan mangrove serta pengelolaan sampah dapat mengurangi dampak buruk perubahan iklim. Hasil ini menguatkan penelitian-penelitian sebelumnya yang menyatakan bahwa banyak kegiatan pengelolaan lingkungan hidup yang dilakukan di negara-negara berkembang dengan melibatkan pemerintah dan perusahaan melalui berbagai kegiatan sosial yang bertujuan untuk mengurangi dampak perubahan iklim [17].

4) TOPIK 4 (EKONOMI PEMBANGUNAN)

Dalam kategori ini, kata yang dominan adalah “ekonomi,” “kemenkeu” (Kementerian Keuangan), “menko” (menteri koordinator), “tumbuh,” “bank,” “imf” (International Monetary Fund), “proyek,” dan lain-lain. Berdasarkan kata-kata tersebut, dapat diartikan bahwa tema pada topik 4 adalah ekonomi

pembangunan. Sebaran kata pada topik 4 dapat dilihat pada Gambar 6(d).

Kategori ini mencakup kegiatan pembangunan yang dilakukan pemerintah di tengah isu perubahan iklim. Kata “ekonomi,” “bank,” dan “kemenkeu” menunjukkan dampak ekonomi dari perubahan iklim. Beberapa kata seperti “pemulihan” dan “tumbuh” menunjukkan upaya pemulihan perekonomian masyarakat, hal ini juga sejalan dengan penelitian sebelumnya mengenai hubungan perubahan iklim dan dampaknya terhadap perekonomian [18].

5) TOPIK 5 (KERJA SAMA INTERNASIONAL)

Beberapa kata yang paling sering muncul dalam kelompok ini adalah “presiden,” “jokowi” (yaitu nama populer Presiden Republik Indonesia saat ini), “global,” “konferensi,” “isu,” “inggris,” “pertemuan,” “forum,” “asean” (Association of South East Asian Nation), “pemanasan,” “temperatur,” dan lain-lain. Berdasarkan kata-kata tersebut, kerja sama atau

konferensi internasional pemanasan global cocok untuk menjelaskan kategori ini. Detail distribusi kata pada topik ini dapat dilihat pada Gambar 6(e).

Model dapat menangkap tema kerja sama internasional dari teks, yang mana tema ini sangat relevan dengan perubahan iklim. Semua negara terkena dampak perubahan iklim, sehingga kerja sama harus dilakukan untuk menghadapi dampaknya secara global. Hasil ini telah ditunjukkan dalam [19], yang menyatakan bahwa tujuan pengurangan emisi tidak akan mungkin tercapai tanpa adanya kerja sama tingkat internasional.

6) TOPIK 6 (KEBIJAKAN/REGULASI)

Beberapa kata teratas yang muncul dalam kategori ini adalah “kebijakan,” “kerja,” “dpr” (Dewan Perwakilan Rakyat), “pemerintah,” “perjanjian,” “nasional,” dan lain-lain. Tema yang relevan untuk kelompok ini adalah kebijakan atau regulasi. Distribusi kata pada tema ini dapat dilihat pada Gambar 6(f).

Kategori ini mencakup pemberitaan peraturan di tingkat daerah/nasional yang dilakukan oleh pemerintah dan DPR. Temuan ini menunjukkan bahwa telah ada upaya pemerintah untuk mengatasi permasalahan perubahan iklim melalui kebijakan dan peraturan untuk mengurangi dampak perubahan iklim. Hasil ini juga sesuai dengan [20] tentang hubungan antara perubahan iklim dan pengambilan kebijakan selama 30 tahun terakhir, yang menunjukkan pentingnya kebijakan dalam mengatasi fenomena perubahan iklim.

7) TOPIK 7 (REHABILITASI)

Pada kategori ini, beberapa kata yang paling sering muncul antara lain “mangrove,” “solusi,” “rehabilitasi,” “restorasi,” dan “makanan.” Berdasarkan hasil tersebut dapat diartikan bahwa tema yang mewakili kategori ini adalah rehabilitasi atau restorasi. Sebaran kata pada tema ini dapat dilihat pada Gambar 6(g).

Kategori ini menggambarkan upaya pencegahan dampak perubahan iklim melalui kegiatan rehabilitasi dan restorasi. Algoritma mengklasifikasikan upaya restorasi ini sebagai upaya yang sangat terkait dengan pelestarian hutan bakau. Hasil ini cukup masuk akal dan selaras dengan penelitian sebelumnya yang menyatakan bahwa konservasi hutan bakau di kepulauan Indonesia harus menjadi prioritas utama dalam mitigasi perubahan iklim [21].

8) TOPIK 8 (BENCANA ALAM)

Kata-kata yang termasuk dalam topik ini adalah “bmkg,” “hujan,” “cuaca,” “karhutla” (merujuk pada kebakaran hutan dan semak), “ekstrem,” “musim,” “waspada,” dan “kering.” Berdasarkan daftar tersebut, dapat diartikan bahwa topik yang mewakili kategori ini adalah fenomena alam atau bencana. Distribusi kata pada topik 8 dapat dilihat pada Gambar 6(h).

Kategori bencana merupakan aspek yang tidak dapat dipisahkan dari fenomena perubahan iklim [22]. Algoritma pemodelan topik dapat menangkap fenomena ini ke dalam satu kategori fenomena alam dan bencana. Fenomena bencana yang paling banyak ditangkap adalah hujan dan kekeringan ekstrem, kebakaran hutan, dan pergantian musim. Kata “bmkg” sering muncul dalam kategori ini karena merupakan lembaga pemerintah yang selalu memberikan informasi kepada masyarakat mengenai bencana dan mitigasinya.

B. ANALISIS TOPIK

Terdapat beberapa temuan menarik berdasarkan hasil pemodelan topik pada subbab sebelumnya. Secara umum, model tersebut dapat menggambarkan isu-isu penting terkait

perubahan iklim seperti energi terbarukan, bencana, kerja sama internasional, kebijakan, dan ekonomi pembangunan. Hal ini menunjukkan bahwa pembahasan isu tersebut di media cukup komprehensif dalam membahas berbagai aspek perubahan iklim.

Terdapat dua kategori utama yang dapat dicermati dari hasil tersebut, yaitu perubahan iklim terkait fenomena alam dan perubahan iklim terkait regulasi atau kebijakan. Pada kategori perubahan iklim terkait fenomena alam, model berhasil menemukan klaster emisi karbon (topik 2); dan fenomena bencana seperti banjir, kekeringan, dan cuaca ekstrem (topik 8). Pada kategori iklim terkait kebijakan, model berhasil menemukan klaster topik seperti energi terbarukan (topik 1), pengelolaan lingkungan (topik 3), ekonomi pembangunan (topik 4), kerjasama internasional (topik 5), kebijakan (topik 6), dan rehabilitasi (topik 7).

Namun, terdapat beberapa topik yang hilang dari klaster tersebut, seperti tema perubahan iklim yang berkaitan dengan pertanian [3], ketahanan pangan [23], kesehatan [24], dan kemiskinan [25]. Ketiadaan topik-topik terakhir juga menjadi keterbatasan tulisan ini. Setidaknya terdapat dua alasan yang mungkin menyebabkan hasil ini. Pertama, karena terbatasnya rentang waktu atau set data yang digunakan dalam penelitian ini. Kedua, mungkin karena sedikitnya frekuensi *headline* dengan tema-tema di atas, menunjukkan bahwa tema-tema tersebut kurang mendapat perhatian serius.

V. KESIMPULAN

Dalam penelitian ini telah dilakukan pemodelan topik isu perubahan iklim berdasarkan *headline* media di Indonesia. Set data dari beberapa media daring dikumpulkan dengan cara *scraping* situs web masing-masing media. Proses pembersihan data dilakukan untuk mempersiapkan data sebelum mengimplementasikan algoritma *clustering*. Kemudian, algoritma LDA diterapkan pada set data untuk memodelkan topik.

Berdasarkan hasil, algoritma berhasil mengategorikan topik utama mengenai perubahan iklim. Topik-topik tersebut adalah energi terbarukan, emisi karbon, pengelolaan lingkungan hidup, ekonomi pembangunan, kerja sama internasional, kebijakan/regulasi, rehabilitasi, dan kebencanaan. Namun, algoritma tersebut gagal mengategorikan aspek penting lainnya seperti pertanian, ketahanan pangan, kesehatan, dan kemiskinan.

Untuk penelitian selanjutnya, set data yang lebih besar dan rentang waktu yang lebih lama akan digunakan agar mendapatkan hasil yang lebih komprehensif. Selanjutnya, beberapa algoritma *clustering* akan digunakan untuk menganalisis set data untuk membandingkan algoritma mana yang memberikan hasil lebih baik dari yang lain.

KONFLIK KEPENTINGAN

Pada penulis menyatakan tidak ada konflik kepentingan.

KONTRIBUSI PENULIS

Konseptualisasi, Anang Kunaefi dan Aris Fanani; metodologi, Anang Kunaefi dan Aris Fanani; validasi, Anang Kunaefi dan Aris Fanani; analisis formal, Aris Fanani; investigasi, Anang Kunaefi; sumber daya, Aris Fanani; kurasi data, Aris Fanani; penulisan—penyusunan draf asli, Anang Kunaefi; penulisan—penelaahan dan penyuntingan, Anang Kunaefi dan Aris Fanani; visualisasi, Anang Kunaefi; pengawasan, Aris Fanani; administrasi proyek, Aris Fanani; akuisisi pendanaan, Aris Fanani.

REFERENSI

- [1] S. Fawzy, A.I. Osman, J. Doran, dan D.W. Rooney, "Strategies for mitigation of climate change: A review," *Environ. Chem. Lett.*, vol. 18, no. 6, hal. 2069–2094, Nov. 2020, doi: 10.1007/s10311-020-01059-w.
- [2] M. Measey, "Indonesia: A vulnerable country in the face of climate change," *Glob. Major. E-J.*, vol. 1, no. 1, hal. 31–45, Jun. 2010.
- [3] N. Herlina dan A. Prasetyorini, "Pengaruh perubahan iklim pada musim tanam dan produktivitas jagung (*Zea mays L.*) di Kabupaten Malang," *J. Ilmu Pertan. Indones.*, vol. 25, no. 1, hal. 118–128, Jan. 2020, doi: 10.18343/jipi.25.1.118.
- [4] Y. Kinose dkk., "Impact assessment of climate change on the major rice cultivar Ciherang in Indonesia," *J. Agric. Meteorol.*, vol. 76, no. 1, hal. 19–28, Jan. 2020, doi: 10.2480/agrmet.D-19-00045.
- [5] K.A. Harvian dan R.J. Yuhan, "Kajian perubahan iklim terhadap ketahanan pangan," *Semin. Nas. Official Statist.*, 2020, hal. 1052–1061, doi: 10.34123/semnasoffstat.v2020i1.593.
- [6] R. Cahyaningsih dkk., "Climate change impact on medicinal plants in Indonesia," *Glob. Ecol. Conserv.*, vol. 30, hal. 1–13, Oct. 2021, doi: 10.1016/j.gecco.2021.e01752.
- [7] T.R. Keller dkk., "News media coverage of climate change in India 1997–2016: Using automated content analysis to assess themes and topics," *Environ. Commun.*, vol. 14, no. 2, hal. 219–235, Feb. 2020, doi: 10.1080/17524032.2019.1643383.
- [8] P. Swarnakar dan A. Modi, "NLP for climate policy: Creating a knowledge platform for holistic and effective climate action," 2021, *ArXiv: 2105.05621*.
- [9] A. Minnich dkk., "ClearView: Data cleaning for online review mining," 2016 *IEEE/ACM Int. Conf. Adv. Soc. Netw. Anal. Min. (ASONAM)*, 2016, hal. 555–558, doi: 10.1109/ASONAM.2016.7752290.
- [10] D.M. Blei, A.Y. Ng, dan M.I. Jordan, "Latent Dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, hal. 993–1022, Jan. 2003.
- [11] A. Kunaefi dan M. Aritsugi, "Characterizing user decision based on argumentative reviews," 2020 *IEEE/ACM Int. Conf. Big Data Comput. Appl. Technol. (BDCAT)*, 2020, hal. 161–170, doi: 10.1109/BDCAT50828.2020.00002.
- [12] H. Li, R. Lin, R. Hong, dan Y. Ge, "Generative models for mining latent aspects and their ratings from short reviews," 2015 *IEEE Int. Conf. Data Min.*, 2015, hal. 241–250, doi: 10.1109/ICDM.2015.28.
- [13] R. Řehůřek dan P. Sojka, "Software framework for topic modelling with large corpora," *Proc. LREC 2010 Workshop New Chall. NLP Framew.*, 2010, hal. 45–50.
- [14] C. Sievert dan K. Shirley, "LDAvis: A method for visualizing and interpreting topics," *Proc. Workshop Interact. Lang. Learn., Vis. Interfaces*, 2014, hal. 63–70, doi: 10.3115/v1/W14-3110.
- [15] A.G. Olabi dan M.A. Abdelkareem, "Renewable energy and climate change," *Renew. Sustain. Energy Rev.*, vol. 158, hal. 1–7, Apr. 2022, doi: 10.1016/j.rser.2022.112111.
- [16] J. Hansen dkk., "Assessing 'dangerous climate change': Required reduction of carbon emissions to protect young people, future generations and nature," *PLoS ONE*, vol. 8, no. 12, hal. 1–26, Des. 2013, doi: 10.1371/journal.pone.0081648.
- [17] J.A.P. de Oliveira dan C.J.C. Jabbour, "Environmental management, climate change, CSR, and governance in clusters of small firms in developing countries: Toward an integrated analytical framework," *Bus. Soc.*, vol. 56, no. 1, hal. 130–151, Jan. 2017, doi: 10.1177/0007650315575470.
- [18] D. Castells-Quintana, M.P. Lopez-Urbe, dan T.K.J. McDermott, "Adaptation to climate change: A review through a development economics lens," *World Dev.*, vol. 104, hal. 183–196, Apr. 2018, doi: 10.1016/j.worlddev.2017.11.016.
- [19] M.M. Ferrari and M.S. Pagliari, "No country is an island. International cooperation and climate change," Banque de France Working Paper, WP #815
- [20] J. Gupta, "A history of international climate change policy," *WIREs Clim. Change*, vol. 1, no. 5, hal. 636–653, Sep. 2010, doi: 10.1002/wcc.67.
- [21] D. Murdiyarto dkk., "The potential of Indonesian mangrove forests for global climate change mitigation," *Nat. Clim. Change*, vol. 5, no. 12, hal. 1089–1092, Des. 2015, doi: 10.1038/nclimate2734.
- [22] J. Jetten dkk., "Responding to climate change disaster," *Eur. Psychol.*, vol. 26, no. 3, hal. 161–171, Jul. 2021, doi: 10.1027/1016-9040/a000432.
- [23] T. Wheeler dan J. von Braun, "Climate change impacts on global food security," *Sci.*, vol. 341, no. 6145, hal. 508–513, Agu. 2013, doi: 10.1126/science.1239402.
- [24] World Health Organization (2021) "COP26 special report on climate change and health: The health argument for climate action," [Online], <https://apps.who.int/iris/bitstream/handle/10665/346168/9789240036727-eng.pdf?sequence=1>, tanggal akses: 16-Agu-2023.
- [25] S. Hallegatte dan J. Rozenberg, "Climate change through a poverty lens," *Nat. Clim. Change*, vol. 7, no. 4, hal. 250–256, Apr. 2017, doi: 10.1038/nclimate3253.