

Social Media Analysis and Topic Modeling: Case Study of Stunting in Indonesia

Analisis Sosial Media dan Pemodelan Topik: Kasus Studi Stunting di Indonesia

Amri Muhaimin^{1*}, Tresna Maulana Fahrudin², Syifa Syarifah Alamiyah³, Heidy Arviani⁴, Ade Kusuma⁵, Allan Ruhui Fatmah Sari⁶, Angela Lisanthoni⁷

^{1,2,6,7}Sains Data, Universitas Pembangunan Nasional Veteran Jawa Timur, Indonesia

^{3,4,5}Ilmu Komunikasi, Universitas Pembangunan Nasional Veteran Jawa Timur, Indonesia

^{1*} amri.muhamin.stat@upnjatim.ac.id,

*: *Penulis korespondensi (corresponding author)*

Informasi Artikel

Received: September 2023

Revised: October 2023

Accepted: October 2023

Published: October 2023

Abstract

Purpose: Stunting is a problem that currently requires special attention in Indonesia. The stunting rate in 2022 will drop to 21.6%, and for the future, the government has set a target of up to 14% in 2024. Rapid technological developments and freedom of expression on the internet produce review text data that can be analyzed for evaluation. This study analyzes the text data of Twitter users' reviews on stunting. The method used is a text-mining approach and topic modeling based on Latent Dirichlet Allocation.

Design/methodology/approach: The methodology used in this study is Latent Dirichlet Allocation. The data was collected from twitter with the keyword "stunting." After, the data was cleaned and then modeled using the Latent Dirichlet Allocation.

Findings/results: The results show that negative sentiment dominates by 60.6%, positive sentiment by 31.5%, and neutral by 7.9%. In addition, this research shows that 'children,' 'decrease,' 'number,' 'prevention,' and 'nutrition' are among the words that often appear on stunting.

Originality/value/state of the art: This study uses the keyword stunting and analyzes it. Social media analytics show that the people of Indonesia are primarily aware of stunting. Also, the Latent Dirichlet Analysis can be used to create the model.

Abstrak

Tujuan: Stunting merupakan permasalahan yang saat ini memerlukan perhatian khusus di Indonesia. Angka stunting

Keywords: Stunting; Sentiment Analysis; Topic Modelling..

Kata kunci: Analisis Sentiment;
Pemodelan Topik; *Stunting*.

pada tahun 2022 akan turun menjadi 21,6%, dan ke depannya pemerintah telah menetapkan target hingga 14% pada tahun 2024. Perkembangan teknologi yang pesat dan kebebasan berekspresi di internet menghasilkan data teks ulasan yang dapat dianalisis untuk dievaluasi. Penelitian ini menganalisis data teks ulasan pengguna Twitter mengenai *stunting*. Metode yang digunakan adalah pendekatan text-mining dan pemodelan topik berdasarkan Latent Dirichlet Allocation.

Perancangan/metode/pendekatan: Metodologi yang digunakan dalam penelitian ini adalah Latent Dirichlet Allocation. Data dikumpulkan dari twitter dengan kata kunci “*stunting*”. Setelah itu, data dibersihkan dan kemudian dimodelkan menggunakan Latent Dirichlet Allocation.

Hasil: Hasilnya menunjukkan sentimen negatif mendominasi sebesar 60,6%, sentimen positif sebesar 31,5%, dan netral sebesar 7,9%. Selain itu, penelitian ini menunjukkan bahwa 'anak-anak', 'penurunan', 'angka', 'pencegahan', dan 'gizi' merupakan kata-kata yang sering muncul pada *stunting*.

Keaslian: Penelitian ini menggunakan kata kunci *stunting* dan menganalisisnya. Analisis media sosial menunjukkan bahwa masyarakat Indonesia pada dasarnya sadar akan *stunting*. Selain itu, Analisis Dirichlet Laten dapat digunakan untuk membuat model.

1. Pendahuluan

Stunting merupakan suatu keadaan anak mengalami gangguan pertumbuhan, sehingga tinggi badan anak tidak normal yang disebabkan masalah kekurangan asupan gizi dalam waktu yang lama [1]. Permasalahan *stunting* di Indonesia masih cukup besar pada bidang kesehatan saat ini. Menurut laporan yang diterbitkan oleh *World Health Organization (WHO)*, sebanyak 22% atau sekitar 149,2 juta anak di dunia yang berusia di bawah lima tahun tercatat mengalami *stunting* pada tahun 2020 [2]. Kedudukan Indonesia terhadap prevalensi *stunting* di dunia berada pada urutan 115 dari 151 negara [3]. Sedangkan di Asia Tenggara, Indonesia menduduki peringkat kedua yang mencapai 31,8% setelah Timor Leste sebanyak 48,8% [4]. Urutan ketiga adalah negara Laos sebanyak 30,2%, keempat Kamboja sebanyak 29,9%, dan kelima diduduki oleh negara Filipina sebanyak 29,9%.

Stunting disebabkan oleh masalah kesehatan, faktor lingkungan dan pelayanan kesehatan yang diterima oleh anak. Faktor genetik tidak memengaruhi terjadinya *stunting* secara signifikan. Kurangnya pemberian gizi pada janin menjadi penyebab terbesar kondisi *stunting* pada anak. Usia 1000 hari pertama kehidupan seorang anak (1000 HPK) merupakan titik awal untuk memberikan simpulan penting pada pertumbuhan jangka panjang [5]. Sehingga jika pola asuh dan pola makan yang tidak efektif dapat meningkatkan peluang terjadinya *stunting*. Gangguan mental dan hipertensi pada Ibu juga memberikan pengaruh terhadap perilaku dan praktik

pemberian gizi pada anak. Keterbatasan akses pelayanan kesehatan dan sanitasi memperparah kondisi *stunting* yang terjadi di Indonesia seperti kekurangan air bersih, tempat jamban yang kurang bersih, dan lain sebagainya [6].

Permasalahan *stunting* menjadi perhatian khusus di Indonesia karena akan memengaruhi kualitas sumber daya manusia baik dalam jangka pendek maupun jangka panjang. *Stunting* dapat mengakibatkan anak memiliki keterlambatan perkembangan fungsi otak, perkembangan motorik dan kekebalan tubuh yang buruk. Hal tersebut, akan mengganggu pertumbuhan, produktivitas dan prestasi pada saat beranjak dewasa. Selain itu, akan muncul berbagai penyakit kronis dan jika serius maka dapat memunculkan resiko kematian yang besar [7]. Upaya pemerintah untuk menekan angka *stunting* yang berjalan saat ini adalah pelaksanaan Strategi Nasional Percepatan Pencegahan Stunting 2018-2024, 1000 PJK dan PMT. Beberapa program yang dilakukan adalah peningkatan asupan gizi dan pelayanan pada ibu hamil dan remaja putri, memantau perkembangan dan pertumbuhan ibu hamil, kampanye atau sosialisasi terkait gizi, pemberian tablet tambah darah, dan lain-lain [8].

Selain itu, pemanfaatan Sistem Pemerintah Berbasis Elektronik (SPBE) sebagai basis data dalam menurunkan kasus *stunting* sudah diterapkan dengan baik di beberapa daerah. Penerapan teknologi digital dan analisis data sangat diperlukan untuk mendukung pelaksanaan program yang ada untuk dapat mengevaluasi secara akurat dan tepat. Namun, seiring dengan perkembangan teknologi digital, data yang diperoleh untuk analisis tidak terbatas pada data yang dimiliki oleh instansi tertentu, tetapi dapat dilakukan penambangan data dari teks ulasan yang hampir setiap hari diunggah pada media sosial. Akan tetapi, beragamnya pembahasan pada teks ulasan atau komentar menyebabkan beragamnya pemaknaan pula. Sehingga diperlukan pemodelan yang mengidentifikasi topik pembahasan yang dominan.

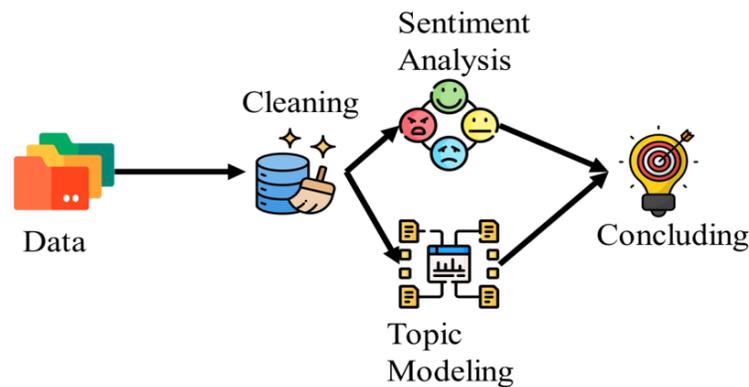
Ulasan yang dimanfaatkan untuk sumber informasi atau data penelitian sudah sering dilakukan oleh peneliti sebelumnya. Pada tahun 2021 oleh Mochamad Yudha Febrianta, dkk., menganalisis ulasan produk *Indie Video Game* di *Steam* berbasis *machine learning* secara otomatis. Metode yang digunakan adalah analisis sentimen menggunakan algoritma *Naïve Bayes* dan pemodelan topik berbasis *Latent Dirichlet Allocation*. Hasil menunjukkan sentimen positif dominan sebesar 69,8% dengan akurasi algoritma 75,45% [9]. Kemudian, terdapat penelitian untuk menentukan topik teks suatu berita dengan menggunakan metode LDA yang dilakukan oleh Sely Karmila, dkk. pada tahun 2022. Penelitian ini berhasil menerapkan sistem penentuan topik berita menggunakan LDA yang menghasilkan tiga topik dengan akurasi sebesar 67% yang relevan [10]. Pada tahun 2023, penelitian untuk mengidentifikasi suatu topik dari kumpulan teks dan pemodelan topik pada ulasan pengguna aplikasi SatuSehat dilakukan oleh Fikri Fahru Roji, dkk..Metode yang digunakan adalah *Support Vector Machine* (SVM) untuk analisis sentimen dan LDA untuk pemodelan topik. Kesimpulan yang didapat mengindikasikan dominasi tanggapan positif dengan 8 topik teridentifikasi perhatian terhadap kesehatan hingga apresiasi terhadap aplikasi. Selain itu, topik negatif yang ditemukan adalah tantangan vaksinasi, masalah akses, dan fungsionalitas aplikasi [11]. Selanjutnya [12] juga pernah melakukan pemodelan topik menggunakan LDA pada data hasil *crawling* aplikasi *shopee* di *playstore*. Didapatkan hasil terbentuk 10 topik yang paling baik setelah melakukan percobaan 20 topik.

Berdasarkan uraian di atas, penelitian ini bertujuan untuk mengetahui sentimen tentang *stunting*. Untuk menentukan pola asuh dan pola makan yang tidak efektif dapat meningkatkan peluang terjadinya *stunting* [13]. Data yang digunakan berasal dari hasil penambangan data tweet dari

twitter. Kemudian data dikelola untuk pemodelan topik menggunakan *Latent Dirichlet Allocation* (LDA) dengan tujuan mendapatkan topik yang terbentuk dari ulasan terkait permasalahan *stunting* di Twitter.

2. Metode/Perancangan

Bagian ini membahas tentang metode-metode yang digunakan dalam menyelesaikan penelitian terkait. Terdapat empat metode yang digunakan dalam penelitian ini, Penambangan Data, Analisis Teks, Analisis Sentimen, Pemodelan Topik, dan Wordcloud. Untuk lebih lengkapnya, proses analisis atau metode dalam penyelesaian masalah ini disajikan pada gambar 1 berikut.



Gambar 1. Diagram alir penelitian

Terdapat beberapa tahap yang dilakukan dalam penelitian ini yang selanjutnya dijelaskan pada masing-masing sub-bab. Mendapatkan data, kemudian melakukan pembersihan pada data. Setelah data bersih, dilakukan analisis sentimen dengan pemodelan topik. Diakhiri dengan penarikan kesimpulan dan pemberian rekomendasi.

2.1. Penambangan Data

Penambangan data pada penelitian ini dilakukan dengan teknik *scrap* dari Twitter menggunakan API berbayar dari Twitter Developer. Fokus *keyword* yang digunakan adalah topik *stunting* pada bulan Agustus 2023.

2.2. Analisis Teks

Analisis teks adalah suatu proses atau teknik membaca dan memahami teks menggunakan sistem komputer untuk mendapatkan informasi tertentu. Berbagai jenis teks dapat diekstraksi untuk mengambil intisari, makna, dan nilai yang terkandung di dalamnya. Metode yang digunakan tidak memiliki batasan tertentu sehingga disesuaikan dengan tujuan dan kebutuhan di dalam penelitian. Analisis teks biasanya digunakan pada konteks sosial, kesehatan, budaya, politik dan lain sebagainya. Dalam melakukan analisis teks, penelitian ini menggunakan metode tertentu yang dapat memberikan pemahaman mendalam terhadap konten yang diekstraksi. Metode yang digunakan melibatkan langkah-langkah di antaranya pengumpulan data, prapemrosesan teks dengan cara membersihkan dan mengorganisasi teks (penghapusan tautan, karakter khusus, dan pengonversian teks menjadi huruf kecil untuk konsistensi), tokenisasi dengan memperlakukan setiap kata sebagai entitas terpisah, dan ekstraksi fitur-fitur yang relevan dari teks untuk mendukung analisis lebih lanjut. Analisis teks juga merupakan suatu

proses atau teknik membaca dan memahami teks menggunakan sistem komputer untuk mendapatkan informasi tertentu. Berbagai jenis teks dapat diekstraksi untuk mengambil intisari, makna, dan nilai yang terkandung di dalamnya. Metode yang digunakan tidak memiliki batasan tertentu sehingga disesuaikan dengan tujuan dan kebutuhan di dalam penelitian. Analisis teks biasanya digunakan pada konteks sosial, kesehatan, budaya, politik dan lain sebagainya.

2.3. Analisis Sentimen

Analisis sentimen sering digunakan untuk mengetahui kecenderungan opini yang merepresentasikan perasaan atau ekspresi yang disampaikan di Internet. Analisis sentimen merupakan proses memahami, mengekstrak, dan memanipulasi data teks secara otomatis yang digunakan untuk menangkap informasi sentimen yang terkandung dalam kalimat opini [14]. Analisis sentimen sering digunakan untuk mengetahui kecenderungan opini yang merepresentasikan perasaan atau ekspresi yang disampaikan di Internet. Metode analisis sentimen yang digunakan dalam penelitian ini adalah dengan memanfaatkan library VADER untuk mengklasifikasikan sentimen positif, netral, dan negatif. Hasil dari analisis sentimen kemudian diinterpretasikan untuk memahami kecenderungan umum opini yang mewakili perasaan atau ekspresi dalam teks. Keakuratan analisis sentimen divalidasi dengan perbandingan hasil manual dengan hasil otomatis. Hal ini penting untuk menilai sejauh mana model dapat diandalkan dalam mengidentifikasi sentimen dengan tepat.

2.4. Wordcloud

Penelitian ini memberikan visualisasi *wordcloud* yang dihasilkan dari pemodelan topik yang telah dilakukan. *Wordcloud* merupakan visualisasi yang menunjukkan jumlah kata yang digunakan pada suatu teks, jika semakin besar yang ditampilkan maka semakin sering kata yang muncul atau digunakan, begitu pun sebaliknya. Visualisasi ini dapat merepresentasikan teks secara visual yang menarik.

2.5. Latent Dirichlet Allocation

Latent Dirichlet Allocation (LDA) adalah sebuah model statistik yang digunakan untuk menggambarkan topik dalam koleksi dokumen teks. LDA adalah salah satu teknik yang paling umum digunakan dalam pemodelan topic untuk analisis teks dan pemahaman isi dokumen. Berikut adalah beberapa poin utama tentang Latent Dirichlet Allocation (LDA):

1. Pemodelan Probabilistik: LDA adalah model probabilistik yang digunakan untuk menjelaskan cara dokumen dibuat. Model ini berdasarkan pada asumsi bahwa setiap dokumen dapat dianggap sebagai campuran dari beberapa topik yang ada dalam korpus teks secara keseluruhan.
2. Topik sebagai Distribusi Kata: LDA menganggap bahwa setiap topik dalam koleksi dokumen adalah distribusi kata-kata. Dengan kata lain, setiap topik didefinisikan oleh sekelompok kata yang sering muncul bersama dalam dokumen.
3. Latent Variables: "Latent" dalam LDA mengacu pada variabel yang tidak diamati langsung. Dalam konteks ini, variabel laten adalah distribusi topik dalam setiap dokumen dan distribusi kata dalam setiap topik.
4. Inferensi: LDA melakukan inferensi untuk menemukan distribusi topik dalam dokumen-dokumen dalam korpus. Ini melibatkan proses untuk memperkirakan seberapa banyak setiap topik hadir dalam setiap dokumen dan seberapa sering kata-kata tertentu muncul dalam setiap topik.

5. Aplikasi: LDA banyak digunakan dalam berbagai aplikasi seperti pengelompokan dokumen, analisis sentimen, dan penggalian informasi. Ini membantu mengungkap struktur topik dalam korpus teks besar yang sulit diidentifikasi secara manual.

Dalam praktiknya, LDA adalah alat yang kuat untuk menggali wawasan dari teks dan dapat membantu dalam mengorganisasi dan memahami konten dokumen secara lebih efisien [15]. Setelah topik terbentuk, topik-topik tersebut dapat dievaluasi dengan menggunakan nilai *coherence*. Nilai tersebut berasal dari perhitungan kesamaan antara topik satu dengan yang lain. Jika kesamaan semakin kecil, berarti topik terbentuk dengan baik atau dalam hal ini adalah topik terbentuk dengan distribusi kata yang berbeda antara topik. Jika kesamaan semakin kecil, maka nilai *coherence* yang dihasilkan semakin besar. Nilai perhitungan tersebut dinamakan *pointwise mutual information* yang disajikan pada persamaan (1) berikut.

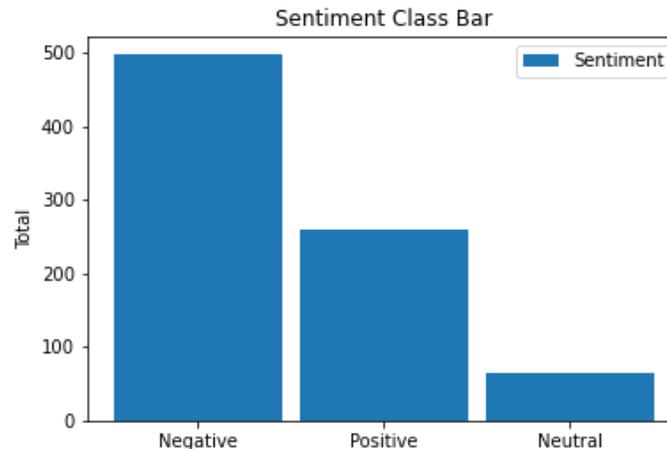
$$pmi(x, y) = \log_2 \frac{p(x, y)}{p(x)p(y)} = \log_2 \frac{p(y|x)}{p(y)} \quad (1)$$

Dengan x dan y adalah urutan dokumen atau urutan topik. Sehingga rumus tersebut adalah membandingkan nilai topik satu dengan yang lain. Yang kemudian nilai PMI tersebut dinormalisasikan sehingga mendapatkan nilai *coherence*.

3. Hasil dan Pembahasan

3.1. Data Ulasan Twitter

Dalam penelitian ini menggunakan data teks ulasan pengguna twitter pada bulan Agustus 2023 mengenai stunting. Setelah dilakukan pembersihan data yang terkumpul, terdapat 822 review yang sudah bersih dan akan masuk ke tahap selanjutnya yaitu proses pemodelan.

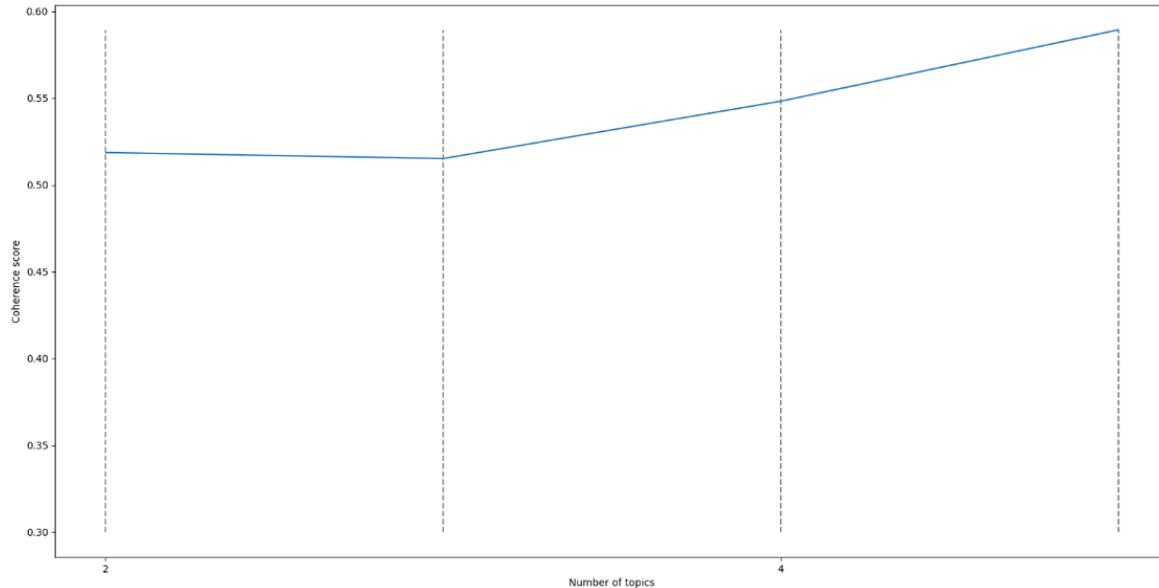


Gambar 2. Sentimen dalam Data Ulasan Twitter

Berdasarkan Gambar 2, terlihat sebaran data ulasan pengguna twitter stunting mempunyai puncak tertinggi pada kelas sentimen negatif sebanyak 498 ulasan. Artinya, ketika membahas topik stunting, pengguna twitter cenderung memberikan ulasan negatif. Namun, total ada 259 ulasan yang masih memberikan respon positif terhadap topik tersebut. Sisanya, sebanyak 65 pengguna twitter memberikan tanggapan netral. Terlihat jelas bahwa jumlah ulasan negatif dua

3.2. Pemodelan Topik dengan LDA

Pemodelan dilakukan dari data teks yang telah diolah dan dibersihkan. Kemudian dilakukan sebuah percobaan untuk menentukan jumlah topik yang optimal yang dapat dibuat dari data yang teks tersebut. Hasil eksperimen disajikan pada Gambar 4.



Gambar 2. Grafik Nilai Kebajikan Topik yang Terbentuk

Percobaan yang dilakukan dengan membentuk model topik sebanyak lima model, kemudian mengukur kebaikan topik tersebut dengan nilai *coherence*. Didapatkan bahwa topik yang optimal adalah lima topik. Nilai *coherence* maksimal yang dihasilkan adalah ketika data di modelkan kedalam 5 topik. Namun perbedaan antara empat topik dan lima topik tidak begitu signifikan. Sehingga dalam hal ini, empat topik dipilih karena akan menghasilkan model yang lebih sederhana dengan nilai kebaikan yang sama dengan lima topik. Model yang terbentuk disajikan dalam Persamaan 1 berikut:

$$\begin{aligned}y_1 &= 0.045_{turun} + 0.028_{rektonoseto} + 0.027_{angka} \\y_2 &= 0.062_{program} + 0.039_{dampak} + 0.027_{bkkbn} \\y_3 &= 0.053_{target} + 0.035_{posyandu} + 0.030_{bantu} \\y_4 &= 0.076_{turun} + 0.044_{rektonoseto} + 0.029_{angka}\end{aligned}\tag{1}$$

Berdasarkan Persamaan (1), jumlah model topik yang terbentuk adalah 4 topik. Masing-masing topik mempunyai kata yang paling mewakili topik tersebut. Misalnya pada y_1 adalah topik model 1, mempunyai kata turun, rektonoseto, dan angka, kemungkinan pada topik tersebut adalah membicarakan cara menurunkan angka stunting. Model 1 dan model 4 mempunyai kata yang sama, tapi nilai kontribusinya berbeda. Sedangkan pada model 2 adalah program, dampak, dan bkkbn. Kemungkinan membahas BKKBN untuk membuat program sehingga meringankan dampak stunting atau menurunkan dampak anak kurang gizi. Kemudian di model 3 kata yang berkontribusi adalah target, posyandu, dan bantu. Pada model ini kemungkinan adalah bagaimana target untuk melakukan penurunan stunting yang dibantu oleh posyandu.

Pembentukan topik didasarkan oleh distribusi kata yang sering muncul pada masing-masing topik model. Kemudian kata-kata itu dirangkum dan dibuat topik sesuai dengan distribusi kata tersebut.

3.3. Rekomendasi

Berdasarkan hasil kajian yang telah dilakukan, dapat diketahui bahwa para ibu rumah tangga sebagian besar masih mengandalkan susu untuk pemenuhan kebutuhan gizi anak-anaknya. Sebagian sentimen yang terbentuk adalah sentimen negatif, mayoritas bersentimen negatif. Hal tersebut dikarenakan mungkin karena tidak terdapat peran pemerintah dalam mengatasi stunting. Meskipun angka stunting di Indonesia lambat laun menurun, namun angka tersebut masih cukup tinggi jika dibandingkan dengan negara maju lainnya. Hal-hal yang dapat dilakukan oleh pemerintah adalah mendirikan aplikasi

4. Kesimpulan dan Saran

Analisis yang dilakukan pada data ulasan pengguna twitter terhadap stunting di bulan Agustus 2023, menunjukkan bahwa sebaran data ulasan mempunyai puncak tertinggi pada kelas sentimen negatif sebanyak 498 ulasan. Sedangkan ulasan positif sebanyak 259 dan ulasan netral sebanyak 65. Sentimen negatif mendominasi topik stunting yang mana terdapat lima kata teratas yang sering muncul pada ulasan, seperti 'anak', 'turun', 'angka', 'cegah' dan 'gizi'. Kemudian, pada visualisasi *wordcloud* terdapat kata-kata yang muncul selain lima belas kata teratas, seperti 'miskin', 'posyandu', 'kerja', dan lain-lain. Hal tersebut menunjukkan bahwa masalah stunting dapat dikaitkan dengan implemmentasi program pemerintah terhadap tindakan preventif, penyediaan gizi, faktor taraf kemiskinan, dan fasilitas akses kesehatan masyarakat yang perlu dianalisis lebih dalam lagi untuk dapat ditingkatkan. Pemodelan topik menggunakan LDA dihasilkan 4 topik dengan nilai coherence sebesar 0.6, masing-masing dari model mempunyai ciri bahasan mengenai topik yang dimodelkan.

Daftar Pustaka

- [1] H. Rahman, M. Rahmah and N. Saribulan, "Upaya Penanganan Stunting di Indonesia Analisis Bibliometrik dan Analisis Konten," *Jurnal Ilmu Pemerintahan Suara Khatulistiwa (JIPSK)*, vol. 8, no. 1, Juni 2023.
- [2] W. H. Organization, *World Health Statistics 2022:Monitoring Health for the SDGs*, World Health Organization, 2022.
- [3] Novrizaldi, "Kementrian Koordinasi Bidang Pembangunan Manusia dan Kebudayaan Republik Indonesia," KEMENKO PMK, 23 Agustus 2021. [Online]. Available: <https://www.kemenkopmk.go.id/menko-pmk-beberkan-kunci-atasi-gizi-buruk-dan-stunting>. [Accessed September 2023].
- [4] T. A. D. B. (ADB), *Asian Development Outlook (ADO) 2020: What Drives Innovation in Asia?*, Asian Development Bank, 2020, p. 396.
- [5] M. I. Panigoro, A. A. Sudirman and D. Modjo, "Upaya Pencegahan dan Penanggulangan Stunting pada Balita di Wilayah Kerja Puskesmas Tilongkabila," *Jurnal Ilmu Kesehatan dan Gizi (JIG)*, vol. 1, no. 1, pp. 47-60, 2023.

- [6] R. D. R. S. Blora, "RSUD Dr. R Soetijono Blora," 15 December 2022. [Online]. Available: <https://rsudblora.blorakab.go.id/2022/12/15/mengenal-stunting-penyebab-hingga-cara-pencegahannya/#:~:text=Penyebabnya%2C%20adalah%20karena%20sang%20ibu,ikut%20memengaruhi%20kondisi%20malnutrisi%20janin.> [Accessed September 2023].
- [7] S. K. Nisa, E. D. Lustiyati and A. Fitriani, "Sanitasi Penyediaan Air Bersih dengan Kejadian Stunting pada Balita," *Jurnal Penelitian dan Pengembangan Kesehatan Masyarakat Indonesia*, vol. 2, no. 1, pp. 17-25, 2021.
- [8] A. P. Yudaa, Z. Septinaa, A. Maharania and Y. Nurdiatami, "Tinjauan Literatur : Perkembangan Program Penanggulangan Stunting di Indonesia," *Jurnal Epidemiologi Kesehatan Indonesia*, vol. 6, no. 2, 2022.
- [9] M. Y. Febrianta, S. Widiyanesti and S. R. R. , "Analisis Ulasan Indie Video Game Lokal pada Steam Menggunakan Analisis Sentimen dan Pemodelan Topik Berbasis Latent Dirichlet Allocation," *Journal of Animation & Games Studies*, vol. 7, no. 2, 2021.
- [10] S. Karmila and V. I. Ardianti, "Metode Latent Dirichlet Allocation untuk Menentukan Topik Teks Suatu Berita," *Jurnal Informatika dan Komputasi*, vol. 16, no. 1, 2022.
- [11] F. F. Roji, N. G. Ginasta, Y. Cahyan, D. Rahayu and D. Ramdani, "Review Analysis of SatuSehat Application Using Support Vector Machine and Latent Dirichlet Allocation Modeling," *JURNAL RISTEC : Research in Information Systems and Technology*, vol. 4, no. 1, 2023.
- [12] Muhaimin, A., . S, S., & Atnanda, P. (2023, November 7). Analisis Topic Modelling pada Ulasan Aplikasi Shopee di PlayStore Menggunakan Latent Dirichlet Allocation (LDA). *PROSIDING SEMINAR NASIONAL SAINS DATA*, 3(1), 122-133. <https://doi.org/https://doi.org/10.33005/senada.v3i1.91>
- [13] Mughni, M., Fahrudin, T., & Kamisutara, M. (2021). Classification of Toddler Nutritional Status Based on Anthropometric Index and Feature Discrimination using Support Vector Machine Hyperparameter Tuning. *International Journal of Computer, Network Security and Information System (IJCONSIST)*, 2(2), 60-65.
- [14] Aqila, J. J. Sihombing, R. I. Sitorus and Arnita, "Implementasi Algoritma Support Vector Machine Untuk Analisis Sentimen Aplikasi OLX di Playstore," *Journal of Informatics and Data Science (J-IDS)*, vol. 1, no. 2, 2022.
- [15] Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77-84.