



Analisis Kepercayaan Masyarakat Tentang Kepolisian Indonesia di Twitter Menggunakan Latent Dirichlet Allocation (LDA)

Bagas Dwi Santosa^{a,1,*}, Nurul Fatimah^{a,2}, Netania Indi Kusumaningtyas^{a,3}, Ulfi Saidata Aesyri^{a,4},
Herdiesel Santoso^{b,5}

^a Program Studi Sistem Informasi Universitas Jenderal Achmad Yani Yogyakarta, Jl. Siliwangi Ringroad Barat, Sleman 55293, Indonesia

^b Program Studi Sistem Informasi STMIK El Rahma, Jalan Sisingamangaraja No. 76 Karangajen, Yogyakarta 55153, Indonesia

¹ bagasdwisantosa87@gmail.com*; ² nurulfatimahnurul31@gmail.com; ³ netania0412@gmail.com; ⁴ ulfiaesyri@gmail.com;

⁵ herdiesel.santoso@stmikrahma.ac.id

* corresponding author

ABSTRACT

ARTICLE INFO

Kepolisian Negara Republik Indonesia merupakan instansi yang bertugas untuk menjaga ketertiban, menerapkan hukum, dan memberikan perlindungan kepada warga. Namun, beberapa kasus melibatkan anggota polisi, membuatnya dibicarakan di sosial media, khususnya Twitter. Hal ini mengakibatkan tagar-tagar terkait polisi menjadi trending. Oleh karena itu, penting untuk mengukur tingkat kepercayaan masyarakat terhadap kepolisian karena berdampak pada stabilitas sosial, keadilan, dan layanan publik. Penelitian ini dilakukan menggunakan metode *Latent Dirichlet Allocation (LDA)*. Data diambil dari periode saat terjadi banyak kasus yang melibatkan oknum kepolisian, yakni dari rentang tanggal 1 Agustus hingga 19 Oktober 2022. Hasil analisis yang diperoleh dibandingkan dengan pidato arahan presiden guna meninjau opini yang diungkapkan oleh presiden dengan pendapat masyarakat dari media sosial Twitter, yang dapat memberikan perspektif yang lebih luas tentang bagaimana lembaga kepolisian dilihat oleh masyarakat. Hasil pengujian koheren diperoleh jumlah topik yang terbentuk sebanyak 5 topik dengan nilai koherensi sebesar 0.4945. Topik-topik tersebut meliputi pembunuhan brigadir oleh anggota polisi, penanganan kurang tepat oleh pihak kepolisian terkait tragedi Kanjuruhan, peran Kapolda Jatim dalam kasus narkoba, perintah Kapolri kepada anak buahnya untuk mengembalikan kepercayaan masyarakat, dan tanggapan masyarakat terhadap kasus yang melibatkan oknum polisi. Hasil dari setiap topik tersebut menunjukkan pentingnya peningkatan kinerja kepolisian, kebijakan organisasi yang tepat, dan penegakan hukum yang tegas untuk memulihkan dan mempertahankan kepercayaan masyarakat terhadap institusi kepolisian. Hal ini sejalan dengan arahan Presiden yang menekankan perlunya memperbaiki keluhan masyarakat, memberikan pelayanan terbaik, solidaritas internal, kebijakan organisasi yang sesuai, dan penegakan hukum yang tegas.

Article history

Received: 7 September 2023

Revised: 12 November 2023

Accepted: 12 November 2023

Keywords

Kepolisian Negara Republik Indonesia

Kepercayaan Publik

Latent Dirichlet Allocation (LDA)

Sosial Media Twitter

Topic Modelling

This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



I. Introduction

Kepolisian adalah lembaga yang bertanggung jawab dalam menegakkan hukum dan memberikan perlindungan kepada masyarakat. Kepolisian menjadi garda terdepan dalam memelihara keamanan dan ketertiban di lingkungan masyarakat dengan menjunjung tinggi prinsip hak asasi manusia. Dalam menjalankan tugasnya, kinerja kepolisian mendapat banyak perhatian dari masyarakat. Terutama dalam era teknologi digital seperti saat ini, informasi dapat mudah didapatkan melalui media sosial. Twitter menjadi salah satu media yang digunakan masyarakat untuk berpendapat. Twitter dianggap cocok sebagai wadah untuk menyampaikan pendapat karena akses yang mudah, tidak adanya batasan jumlah pengikut, keterbatasan karakter sampai 280 huruf yang memungkinkan masyarakat dapat menyampaikan pendapatnya secara singkat, padat, dan jelas [1].

Beberapa topik yang sedang hangat diperbincangkan di media sosial Twitter seperti sejumlah kasus yang melibatkan anggota polisi, kasus penembakan antar polisi, keterlibatan polisi dalam perdagangan narkoba, dan beragam kasus lainnya. Kepolisian yang seharusnya bertugas sebagai penegak hukum dalam menangani kasus-kasus kejahatan, namun perbuatannya bertolak belakang dengan tugas yang harus dijalankannya. Hal ini berpengaruh pada tingkat kepercayaan masyarakat terhadap kepolisian Republik Indonesia. Karena hal tersebut, diperlukan analisis untuk mengevaluasi tingkat kepercayaan masyarakat terhadap pihak kepolisian, langkah ini bertujuan untuk memperoleh kesimpulan yang dapat digunakan sebagai dasar perbaikan bagi pihak kepolisian. Analisis mengenai tingkat kepercayaan masyarakat terhadap kepolisian menggunakan metode *Latent Dirichlet Allocation* (LDA) guna menghasilkan representasi dan hubungan dalam setiap topik yang dihasilkan [2].

Pemodelan topik adalah metode otomatis untuk mengidentifikasi topik dalam teks dan menemukan pola tersembunyi didalam korpus teks [3]. *Latent Dirichlet Allocation* (LDA) merupakan permodelan probabilistik yang berguna untuk mengidentifikasi topik dalam kumpulan data [4].

Asumsi LDA dalam pemodelan topik sebagai berikut:

- Dokumen yang memiliki topik sama menggunakan kelompok kata yang mirip.
- Topik laten dapat diidentifikasi dengan mencari kelompok kata yang sering muncul bersamaan dalam dokumen di keseluruhan korpus.
- Distribusi probabilitas dalam topik laten menandakan berisi banyak sebaran kata dari topik tertentu dalam dokumen.

LDA dipilih dalam penelitian ini karena memungkinkan identifikasi topik-topik tersembunyi dalam tweet dengan jumlah yang besar, pemetaan isu-isu utama yang dibahas dan mengekstraksi informasi relevan mengenai pandangan masyarakat terhadap kepolisian. Keunggulan LDA dalam pemrosesan data teks berskala besar yaitu mampu untuk mengekstraksi pola dari volume besar tweet di Twitter. Kombinasi LDA dengan analisis sentimen dan visualisasi topik dapat memberikan wawasan yang lebih mendalam mengenai sentimen dan pemahaman pandangan masyarakat terhadap Kepolisian Indonesia, sehingga memungkinkan interpretasi yang lebih menyeluruh terhadap hasil analisis.

LDA digunakan oleh [5] untuk menentukan topik politik pada Twitter. Selain itu, LDA juga digunakan [6] untuk mengetahui gambaran umum dari perbincangan publik tentang omnibus law. Penelitian yang dilakukan oleh [7] menerapkan metode LDA pada pemberitaan portal berita online pada masa PSBB pertama. Penting untuk membandingkan hasil analisis LDA terhadap opini masyarakat di Twitter dengan pidato presiden terkait Kapolri untuk memberikan gambaran menyeluruh mengenai pandangan masyarakat dan arah kebijakan resmi pemerintah. Hal ini membantu untuk menilai kesesuaian antara prioritas dan isu-isu yang diperhatikan oleh pemerintah dengan kebutuhan dan harapan masyarakat mengenai kepercayaan terhadap Kepolisian Indonesia. Dengan membandingkan opini publik dengan pandangan resmi pemerintah, dapat teridentifikasi

perbedaan, keselarasan, dan hal-hal yang memerlukan perhatian lebih lanjut dalam upaya meningkatkan kepercayaan masyarakat terhadap lembaga kepolisian.

2. Metode Penelitian

2.1. Dataset

Dataset yang diambil adalah data dari platform media sosial Twitter yang menggunakan tagar:

- | | |
|-----------------------|---------------------------|
| 1. Percumalaporpolisi | 7. teddyminahasa |
| 2. Aparatbobrok | 8. oknumpolisi |
| 3. Aparatmiskinmoral | 9. aparatmiskinmoral |
| 4. Kapolri | 10. polisisampahperadaban |
| 5. Kapolda | 11. polisipembunuh |
| 6. Ferdysambo | 12. percumaadapolisi |

Tagar tersebut digunakan karena menjadi *trending* Twitter dan mencakup berbagai pandangan masyarakat terkait Kapolri. Pengambilan data menggunakan metode *web scrapping*, dengan rentang tanggal 1 Agustus hingga 19 Oktober 2022. Data *tweet* hasil *scraping* meliputi username, tanggal posting, *tweet* dan *url tweet*. Selain itu, dilakukan pengambilan data manual pada pidato Presiden yang disampaikan pada 14 Oktober 2022 berkaitan dengan instruksi atau petunjuk kepada Kepolisian Republik Indonesia (Polri). Analisis tersebut digunakan untuk membandingkan opini yang diungkapkan oleh presiden dengan pendapat masyarakat dari media sosial Twitter, yang dapat memberikan perspektif yang lebih luas dan mendalam tentang bagaimana lembaga kepolisian dilihat dan diresapi oleh masyarakat, serta sejauh mana hal tersebut sejalan dengan visi pemerintah.

2.2. Algoritma Latent Dirichelet Allocation (LDA)

Metode *Latent Dirichelet Allocation* (LDA) adalah salah satu metode dalam *unsupervised learning* dan model probabilitas yang menggambarkan dokumen sebagai kombinasi acak dari topik laten, dimana setiap topik memiliki distribusi kata yang mewakili ciri khasnya [8]. *LDA* mengasumsikan bahwa dokumen direpresentasikan sebagai "*bag-of-words*" atau kumpulan kata [9]. Prinsip kerja LDA ditunjukkan pada Fig 1.

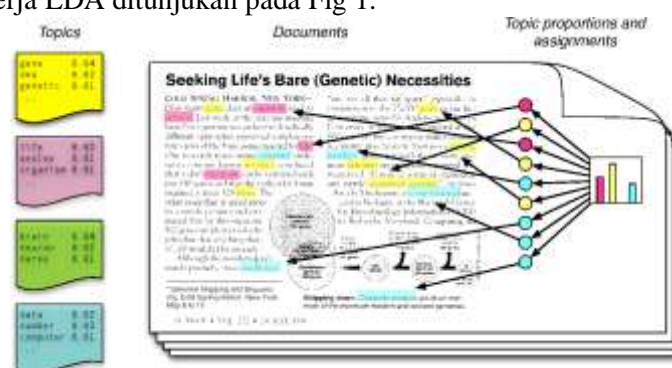


Fig. 1. Prinsip kerja LDA menurut Blei [10]

Penjelasan dari prinsip kerja LDA menjelaskan proses generatif untuk setiap dokumen w dalam kumpulan dokumen D :

- 1) Tetapkan $N \sim \text{Poisson}(x)$,
- 2) Tetapkan $\theta \sim \text{Dir}(a)$,
- 3) Untuk tiap N kata w_n ,
 - a. Tetapkan Topik $z_n \sim \text{Multinomial}(\theta)$,

- b. Tetapkan sebuah kata w_n dari $p(w_n | z_n, \beta)$.

Dalam pemodelan teks, probabilitas topik memberikan representasi yang jelas terhadap suatu dokumen. Misalnya, kita mempunyai suatu dokumen dengan 3 atau 4 topik yang berbeda. LDA akan menganalisis setiap kata dalam dokumen secara random, kemudian menghitung skor untuk masing-masing kata berdasarkan probabilitas kemunculan dalam dokumen yang ditetapkan. LDA selanjutnya akan mengaitkan kata yang identik ke topik-topik lain dan mengevaluasi skornya. Tujuan pemodelan topik adalah mengidentifikasi otomatis topik-topik dari sekumpulan dokumen. Struktur tersembunyi dari data yang diolah berupa topik, distribusi topik disetiap dokumen, serta penentuan untuk setiap kata didalam setiap dokumen.

2.3. Desain dan Implementasi Data

Desain dan implementasi data dapat dilihat pada Fig 2.



Fig. 2. Desain dan Implementasi Data

Desain dan implementasi data meliputi :

- *Scrapping Data Twitter*
Tahap pertama dalam penelitian, mengambil data dari sosial media Twitter menggunakan metode *scrapping* dengan bahasa pemrograman Python dan *Tools* Google Colaboratory.
- *Preprocessing*
Tahap kedua melakukan pembersihan data dari hasil pengambilan data di sosial media Twitter, tahap ini adalah *cleaning text* yang meliputi pembersihan data *null*, data duplikasi, *text lower*, membersihkan angka, tanda hubung dan sebagainya. Tahap selanjutnya *lemmatizing*. *Lemmatizing* adalah langkah yang digunakan untuk mereduksi kata ke bentuk dasar. Tahap terakhir adalah *stopword removal* yang bertujuan untuk menghapus kata-kata penghubung, kata depan, dan kata ganti orang yang tidak memiliki arti dalam dalam suatu konteks[11].
- Pemodelan Topik dengan LDA
Tahap permodelan dilakukan untuk mengidentifikasi topik dalam teks, dalam penelitian ini permodelan menggunakan metode *Latent Dirichlet Allocation* (LDA).
- Visualisasi Data
Tahap Visualisasi digunakan untuk menampilkan hasil permodelan menggunakan metode *Latent Dirichlet Allocation* (LDA). Tahap visualisasi yang dipakai dalam penelitian ini adalah *wordcloud* dan *PyLDAviz*. *Wordcloud* merupakan adalah visualisasi memunculkan kata-kata yang menekankan pada frekuensi kemunculan kata dalam teks [12]. *PyLDAviz* adalah salah satu *library* yang berguna untuk mengekstrak informasi dari hasil permodelan LDA dengan cara menampilkan informasi visual melalui antarmuka web yang interaktif [13].
- Analisis Data
Tahap terakhir adalah menganalisis topik-topik yang dihasilkan dari permodelan LDA, dan membandingkan apakah topik-topik tersebut berhubungan dengan pidato Presiden.

3. Hasil dan Pembahasan

3.1. Data Collecting

Pengambilan data *tweet* pada sosial media Twitter dilakukan menggunakan metode *web scrapping*. *Web scrapping* adalah metode untuk mendapatkan informasi dari *website* secara otomatis tanpa harus menyalinnya secara manual [14]. Data *tweet* yang diambil berisikan username, tanggal posting, isi *tweet* dan *url tweet* tersebut.

Pengambilan data juga dilakukan secara manual untuk mendapatkan topik dari Pidato Presiden pada 14 Oktober 2022 dengan isi pidato yaitu Polri berhasil mereda pandemi covid-19 melalui vaksin yang tersebar sehingga ekonomi tumbuh 5,44%, dan indeks kepercayaan

masyarakat terhadap Polri ada di puncak teratas pada saat itu. Tetapi setelah ada peristiwa FS, kepercayaan masyarakat terhadap Polri turun, yang awalnya pada November 2021 kepercayaan publik terhadap Polri masih 80,2%, tapi Agustus 2022 turun menjadi 54%. Maka Presiden Jokowi memberikan arahan pada Polri seperti pada Tabel 1.

Table 1. Arahan Presiden [15]

No	Arahan
1.	Mengatasi masalah keluhan masyarakat terkait pungli, mencari-cari kesalahan dan gaya hidup mewah.
2.	Polri perlu menegaskan kepada anggotanya untuk memberikan pelayanan terbaik kepada masyarakat agar tercipta rasa aman dan nyaman.
3.	Menunjukkan hubungan yang solid antara internal Polri maupun dengan TNI.
4.	Penerapan kebijakan organisasi sesuai SOP sehingga mencapai kesamaan visi.
5.	Pemerintah dan Polri perlu lebih tegas dalam penegakan hukum agar tidak dianggap lemah

3.2. Data Preparation

Setelah tahap pengambilan data, tahap berikutnya adalah *preprocessing* data. Data yang telah terkumpul perlu dicek untuk memeriksa apakah ada data duplikat atau bernilai kosong (*null*). Tahap pertama dalam *preprocessing* adalah *cleaning text*, langkah *cleaning text* melibatkan normalisasi teks dengan mengonversi semua huruf menjadi huruf kecil, diikuti dengan menghapus angka, tanda hubung, dan karakter simbol lainnya. Tahap berikutnya adalah *lemmatizing*, yang berguna untuk menghapus kata imbuhan disetiap kata yang terdapat imbuhan. Tahap terakhir dalam *preprocessing* adalah *stopword removal*, yang berguna menghapus kata tidak bermakna dalam tweet. Tujuan dari tahapan ini adalah untuk memperoleh dataset yang bersih untuk mendapatkan kualitas analisis topik yang optimal. Jumlah data yang diambil dari Twitter berjumlah 47660 *tweet* dan setelah melalui tahap *preprocessing* menjadi 20361 *tweet*. Contoh tahap *preprocessing* dapat dilihat pada Table 2.

Table 2. Tahap *Preprocessing*

Topic	Word
Data Asli	2022-10-18 12:54:19+00:00,1582354388894781442,"Kapolri akan evaluasi kapolda yang tak mampu kembalikan kepercayaan public
<i>Cleaning text</i>	kapolri akan evaluasi kapolda yang tak mampu kembalikan kepercayaan public
<i>Lemmatizing</i>	Kapolri akan evaluasi kapolda yang tak mampu kembali percaya public
<i>Stopword Removal</i>	Kapolri evaluasi kapolda tidak mampu kembali percaya public

3.3. Pemodelan Topik

3.3.1. Token dan Phrase Model

Tahap pertama dalam pemodelan topik adalah membentuk token dari *tweet* yang sudah dibersihkan ditahap *preprocessing*. *Tweet* yang sebelumnya berbentuk kalimat akan dipisah menjadi kata per kata. *Phrase modelling* berguna untuk membuat model *bigram* dan *trigram*. Tujuan pembentukan *bigram* dan *trigram* adalah untuk mengekstrak arti kata yang terdiri dari dua kata dan tiga kata. Langkah ini berguna untuk mengurangi pemotongan kata yang dapat menyebabkan kehilangan makna dalam kata tersebut [16]. Contohnya penerapan *Bigram* dan *Trigram* dilihat pada Table 3.

Table 3. Contoh Pembentukan Token dan Parse model

Tahap	Hasil
Teks	pulih yakin masyarakat terhadap kepolisian
Bigram	["pulih yakin", "yakin masyarakat", "masyarakat terhadap", "terhadap kepolisian"]
Trigram	["pulih yakin masyarakat", "masyarakat terhadap kepolisian"]

3.3.2. Dictionary dan Corpus

Tahap selanjutnya dalam pemodelan topik adalah pembentukan *dictionary* dan *corpus*. *Dictionary* berisi kumpulan seluruh token, sedangkan *corpus* berisikan token yang telah dikonversi menjadi angka sebagai identitas token dan mencakup jumlah kemunculan setiap token didalam setiap dokumen [17]. Contoh penerapannya seperti pada Table 4.

Table 4. Contoh Pembentukan Dictionary dan Corpus

Tahap	Hasil
Token	['tuntut', 'akan', 'percaya', 'publik', 'kapolri', 'akan', 'evaluasi']
Corpus	[(1, 1), (2, 2), (3, 1), (4, 1), (5, 1), (6, 1)]

3.3.3. Menentukan Nilai Koherensi

Tahap berikutnya adalah mengevaluasi nilai koherensi. Tujuan mengevaluasi nilai koherensi adalah untuk menentukan jumlah topik yang akan digunakan, dengan mencari nilai tertinggi dalam iterasi nilai koherensi. Jumlah iterasi yang digunakan adalah 10 dan hasilnya dapat dilihat pada Table 5.

Table 5. Nilai Koherensi Setiap Topik

Jumlah topik	Nilai Koherensi
1	0.4769
2	0.4496
3	0.3582
4	0.3830
5	0.4945
6	0.3767
7	0.4127
8	0.4756
9	0.4134

Berdasarkan hasil iterasi nilai koherensi pada Table 5, menentukan jumlah topik yang optimal dapat menggunakan nilai koherensi tertinggi yaitu 5 topik dengan nilai koherensi sebesar 0.4945.

3.3.4. Visualisasi Pemodelan Topik

Setelah melakukan pemodelan topik dengan menentukan jumlah topik menggunakan nilai koherensi, perlu adanya pengamatan terhadap pola pembentukan pada masing-masing jumlah topik dengan menggunakan visualisasi *pyLDAviz* [18]. Hasil visualisasi ini berbentuk peta jarak yang menyatakan persebaran topik, hasil dapat dilihat pada Fig 3.

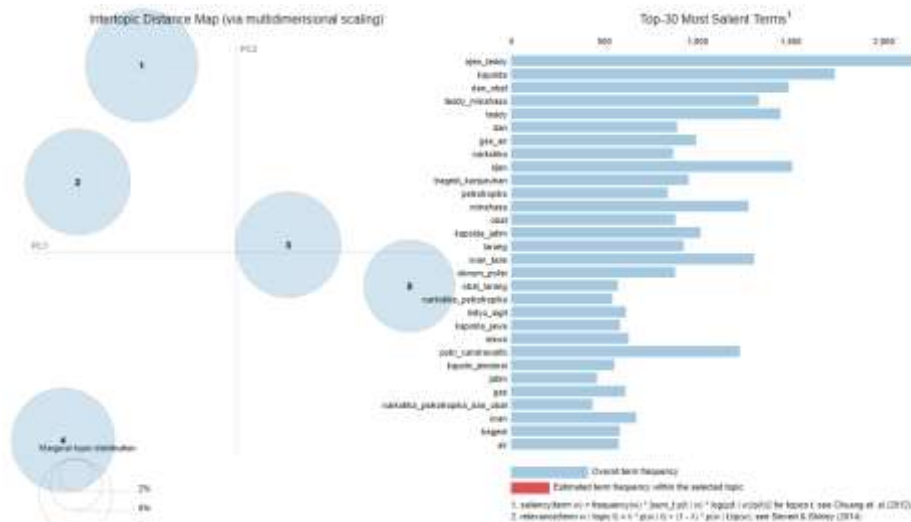


Fig. 3. Visualisasi Distribusi Topik

Dari yang hasil visualisasi distribusi topik dengan jumlah topik 5 menunjukkan adanya tidak adanya topik yang saling beririsan. Dari hasil tersebut, pemilihan jumlah topik bernilai 5 cukup optimal karena tidak ada topik yang beririsan dan persebaran cukup merata. Selain itu, 5 topik dapat memberikan representasi yang cukup baik dalam setiap topik. Hasil setiap topik dapat dilihat pada Table 6.

Table 6. Hasil Setiap Topik

Topik	Kata
1	0.015*"polisi" + 0.015*"bunuh" + 0.012*"siswa" + 0.012*"brigadir" + 0.012*"hukum" + 0.011*"bunuh_brigadir" + 0.009*"piscok" + 0.009*"komnas_ham" + 0.008*"rizky_billard" + 0.007*"putri_candrawathi"
2	0.023*"iwan_bule" + 0.021*"tragedi_kanjuruhan" + 0.018*"putri_candrawathi" + 0.017*"polisi" + 0.012*"tragedi" + 0.012*"kanjuruhan" + 0.011*"bule" + 0.010*"jokowi" + 0.010*"cebong" + 0.009*"cebong_ollie"
3	0.043*"dan_obat" + 0.037*"teddy_minahasa" + 0.026*"dan" + 0.026*"teddy" + 0.025*"narkotika" + 0.025*"obat" + 0.024*"psikotropika" + 0.024*"kapolda_jatim" + 0.024*"minahasa" + 0.023*"larang"
4	0.049*"irjen_teddy" + 0.027*"irjen" + 0.022*"gas_air" + 0.020*"kapolri" + 0.015*"polisi" + 0.014*"listyo_sigit" + 0.013*"teddy" + 0.013*"gas" + 0.012*"kapolri_jenderal" + 0.012*"mata"
5	0.040*"kapolda" + 0.018*"oknum_polisi" + 0.014*"kapolda_jawa" + 0.012*"iwan" + 0.010*"polisi" + 0.009*"oknum" + 0.009*"polda_metro" + 0.008*"the" + 0.007*"jokowi" + 0.006*"timur"

Dari hasil 5 topik sebelumnya, dapat menentukan topik untuk masing-masing *tweet* yang memungkinkan untuk memberikan klasifikasi topik spesifik bagi setiap *tweet*, serta memberikan pemahaman mendalam terhadap topik dalam setiap *tweet*. Berikut 10 sample dominan topik masing-masing *tweet* dapat dilihat pada Fig 4.

Document_No	Dominant_Topic	Topic_Perc_Contrib	Keywords	Text
0	0	2.0	dan, obat, teddy, minahasa, dan, teddy, narkoba	[jmemeramin, pilot, neh, star, wars, laku, bajat
1	1	2.0	dan, obat, teddy, minahasa, dan, teddy, narkoba	[bandar, judi, bandar, narkoba, psikotropika
2	2	2.0	dan, obat, teddy, minahasa, dan, teddy, narkoba	[dah, reset, polisi, anggota, kriminal, laku
3	3	3.0	ijen, teddy, ijen, gas, air, kapoli, polisi	[ihat, esk, lpk, luggang, politik, hri, b
4	4	3.0	ijen, teddy, ijen, gas, air, kapoli, polisi	[astaghfirullah, adzim, bajat, seh, mrk, lakun]
5	5	2.0	dan, obat, teddy, minahasa, dan, teddy, narkoba	[what, the, fuck, what, the]
6	6	0.0	polisi, bunuh, siswa, brigadir, hukum, bunuh, b	[piscok]
7	7	1.0	awan, bule, tragedi, kanjuruhan, putri, candrawati	[astaghfirullah, adzim]
8	8	1.0	awan, bule, tragedi, kanjuruhan, putri, candrawati	[bayak, manusia, laku, perempuan]
9	9	1.0	awan, bule, tragedi, kanjuruhan, putri, candrawati	[blokir, elite, aman, sultudah, bax, sok, sok


Fig. 4. Sample Dominant Topik Setiap *Tweet*

Berdasarkan hasil dari dominan topik dalam *tweet*, dapat dilihat pada kolom ‘Dominant_Topic’ menghasilkan topik yang sesuai dengan *tweet* dan dapat dilihat nilai dominan dalam kolom ‘Topic_Perc_Contrib’.

3.3.5. Analisis


Analisis hasil pemodelan topik digunakan untuk menghasilkan kesimpulan makna dari setiap hasil dalam 5 topik yang sudah ditentukan sebelumnya. Topik yang sebelumnya belum terdapat makna dilakukan analisis mencari distribusi kata-kata yang terkait dengan setiap topik. Distribusi kata yang ditemukan menjadi acuan dalam memberikan interpretasi dari setiap topik. Hasil interpretasi tersebut akan dibandingkan dengan hasil analisis pidato yang disampaikan Presiden terkait arahan untuk Kepolisian Republik Indonesia (Polri).

Table 7. Hasil Topik 1

Keyword	Wordcloud
Polisi, bunuh, siswa, brigadir, hukum, bunuh_brigadir, piscok, komnas_ham, rizky_billard, putri_candrawathi	


Berdasarkan Tabel 7 dapat ditemukan keterkaitan antara *keyword* seperti polisi, bunuh, brigadir, putri_candrawathi. Dari hasil *keyword* tersebut dapat ditrepretasikan makna dari topik 1 “Kepercayaan masyarakat terkait insiden pembunuhan brigadir yang melibatkan seorang anggota Polisi”. Interpretasi tersebut menyiratkan adanya isu kepercayaan masyarakat terhadap institusi kepolisian, khususnya terkait kasus pembunuhan yang melibatkan seorang brigadir polisi dan Putri Candrawathi. Keterkaitan topik ini dengan arahan Presiden yang tercantum dalam Tabel 1, khususnya nomor 4 dan 5 yang berkaitan dengan kebijakan organisasi dan penegakan hukum, menunjukkan relevansi isu tersebut dengan arah kebijakan dan penegakan hukum yang ditekankan oleh pemerintah.

Table 8. Hasil Topik 2

Sebaran Kata	Wordcloud
awan, bule, tragedi, kanjuruhan, putri, candrawathi, polisi, tragedi, kanjuruhan, jokowi, cebong, cebong_ollie	


Berdasarkan Tabel 8 dapat ditemukan keterkaitan antara *keyword* seperti polisi, tragedi_kanjuruhan, usut_tuntas, dan korban. Hasil tersebut menggambarkan topik "Kepercayaan masyarakat terkait tragedi Kanjuruhan yang disebabkan oleh penanganan kurang tepat oleh pihak kepolisian". Interpretasi ini menekankan kepercayaan masyarakat terhadap penanganan suatu insiden, dalam hal ini tragedi Kanjuruhan, yang diduga diakibatkan oleh ketidakmampuan polisi dalam menangani situasi tersebut. Topik ini tidak berkaitan langsung berkaitan dengan arahan yang diberikan oleh Presiden sesuai Tabel 1. Keterkaitan ini menunjukkan bahwa topik tersebut dapat digunakan sebagai bahan evaluasi bagi Kepolisian Republik Indonesia (Polri) untuk memperbaiki kinerja dan mendapatkan kembali kepercayaan masyarakat.

Table. 9. Hasil Topik 3

Sebaran Kata	Wordcloud
dan_obat, teddy_minahasa, dan, teddy, narkotika, obat, psikotropika, polda_jatim, minahasa, larang	

Berdasarkan Tabel 9 dapat ditemukan keterkaitan antara *keyword* seperti kapolda_jatim, narkotika, dan teddy_minahasa. Hasil tersebut dapat menginterpretasikan tentang "Kepercayaan masyarakat terkait keterlibatan Kapolda Jawa Timur dalam kasus penyalahgunaan narkoba". Interpretasi ini menyoroti peran Kapolda Jatim dalam situasi yang melibatkan peredaran narkoba, yang kemungkinan mempengaruhi kepercayaan masyarakat terhadap Polri di wilayah tersebut. Hal ini berkaitan dengan arahan Presiden yang dinyatakan dalam Tabel 1, terutama dalam nomor 4 dan 5 tentang kebijakan organisasi dan penegakan hukum yaitu perlunya penegakan hukum yang tegas dalam Polri. Meskipun topik ini spesifik terkait dengan kasus salah satu Kapolda Jatim, hal ini sejalan dengan upaya memperbaiki kebijakan organisasi dan penegakan hukum yang tegas untuk menjaga kepercayaan masyarakat terhadap kepolisian.

Table. 10. Hasil Topik 4

Sebaran Kata	Wordcloud
irjen_teddy, irjen, gas_air, kapolri, polisi, listyo_sigit, teddy, gas, kapolri_jenderal, mata	

Berdasarkan Tabel 10 dapat ditemukan keterkaitan antara *keyword* seperti kapolri, listyo_sigit, anak_buah, percaya. Keyword yang saling berkaitan tersebut dapat diinterpretasikan makna topik 4 adalah "Instruksi yang diberikan oleh Kapolri Listyo Sigit kepada anggota bawahannya untuk memulihkan kepercayaan masyarakat terhadap institusi kepolisian". Interpretasi ini menunjukkan adanya intruksi yang diberikan oleh Kapolri kepada bawahannya untuk memulihkan kepercayaan masyarakat terhadap institusi Polri. Keterkaitan topik 4 terhadap arahan Presiden tercantum dalam Tabel 1 nomor 2 dengan arahan untuk memberikan pelayanan terbaik kepada masyarakat, menunjukkan bahwa Kapolri sedang mendorong anak buahnya untuk meningkatkan layanan dan kualitas pelayanan kepada masyarakat sehingga dapat memulihkan kembali kepercayaan masyarakat terhadap kepolisian.

- [3] M. A. N. Febriansyach, F. Rashif, G. I. P. Nirvana, and N. A. Rakhmawati, "Implementasi LDA untuk Pengelompokan Topik Tweet Akun Bot Twitter bertagar# covid-19," *CogITO Smart Journal*, vol. 7, no. 1, pp. 170–181, 2021.
- [4] D. Z. T. Kannitha, M. Mustafid, and P. Kartikasari, "Pemodelan Topik Pada Keluhan Pelanggan Menggunakan Algoritma Latent Dirichlet Allocation Dalam Media Sosial Twitter," *Jurnal Gaussian*, vol. 11, no. 2, pp. 266–277, 2022.
- [5] K. H. Nasution, B. P. Adhi, and others, "SISTEM DETEKSI TOPIK POLITIK PADA TWITTER MENGGUNAKAN ALGORITMA LATENT DIRICHLET ALLOCATION," *PINTER: Jurnal Pendidikan Teknik Informatika dan Komputer*, vol. 5, no. 1, pp. 76–83, 2021.
- [6] M. L. C. Chilmi, "Latent dirichlet allocation lda untuk mengetahui topik pembicaraan warganet twitter tentang omnibus law," Fakultas Sains Dan Teknologi UIN Syarif Hidayatullah Jakarta, 2021.
- [7] W. Wahyudin, "APLIKASI TOPIC MODELING PADA PEMBERITAAN PORTAL BERITA ONLINE SELAMA MASA PSBB PERTAMA," in *Seminar Nasional Official Statistics*, 2020, pp. 309–318.
- [8] Y. Kalepalli, S. Tasneem, P. D. P. Teja, and S. Manne, "Effective comparison of LDA with LSA for topic modelling," in *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, 2020, pp. 1245–1250.
- [9] S. Qomariyah, N. Iriawan, and K. Fithriasari, "Topic modeling twitter data using latent dirichlet allocation and latent semantic analysis," in *AIP conference proceedings*, 2019.
- [10] D. M. Blei, "Probabilistic topic models," *Commun ACM*, vol. 55, no. 4, pp. 77–84, 2012.
- [11] K. S. Nugroho, "Dasar Text Preprocessing dengan Python," *Medium.com*, 2019.
- [12] K. Septiani, "Perbandingan Analisis Sentimen Terhadap Pembayaran Digital 'Go-Pay' Dan 'Ovo' Di Media Sosial Twitter Menggunakan Metode Naive Bayes Dan Word Cloud," *Jurnal Ilmu Data*, vol. 2, no. 10, 2022.
- [13] P. Patmawati and M. Yusuf, "Analisis Topik Modelling Terhadap Penggunaan Sosial Media Twitter oleh Pejabat Negara," *Building of Informatics, Technology and Science (BITS)*, vol. 3, no. 3, pp. 122–129, 2021.
- [14] S. Satriajati, S. B. Panuntun, and S. Pramana, "Implementasi web scraping dalam pengumpulan berita kriminal pada masa pandemi COVID-19," in *Seminar Nasional Official Statistics*, 2020, pp. 300–308.
- [15] Sekretariat Presiden, "Pengarahan Presiden Jokowi kepada Pati Mabes Polri, Kapolda, dan Kapolres Se Indonesia," 15 Oktober, Oct. 2022. Accessed: Oct. 18, 2022. [Online]. Available: https://youtu.be/5paYSv_VDtg
- [16] Y. Matira, I. Setiawan, and others, "Pemodelan Topik pada Judul Berita Online Detikcom Menggunakan Latent Dirichlet Allocation," *ESTIMASI: Journal of Statistics and Its Application*, pp. 53–63, 2023.
- [17] F. H. Rachman, *Komputasi Bahasa Alami*. Media Nusa Creative (MNC Publishing), 2020.
- [18] E. Anggraini, "Latent Dirichlet Allocation Untuk Pemodelan Topik Abstrak Dokumen Skripsi," Universitas Islam Indonesia, 2020.