

# Pemodelan Topik Keluhan Masyarakat Pasca Pandemi Menggunakan Metode *Latent Dirichlet Allocation (LDA)*

(Studi kasus: Kabupaten Sleman Tahun 2022)

Ika Lailia Nur Rohmatun Nazila<sup>1</sup>, Dina Tri Utari<sup>2</sup>

Universitas Islam Indonesia<sup>1,2</sup>

20611106@students.uii.ac.id

Transparansi informasi publik dan reformasi birokrasi yang dicanangkan oleh pemerintah untuk mewujudkan *clean government* dan *good government* dapat diimplementasikan dengan melibatkan partisipasi masyarakat. Pengaduan, kritik, dan saran dari masyarakat kepada pemerintah terkait pelayanan publik yang tidak standar, menyimpang atau tidak teratur perlu diperhatikan untuk evaluasi di masa depan. Seiring perkembangan teknologi, aspirasi masyarakat dapat tersampaikan melalui alat pengaduan publik yang telah dikembangkan oleh pemerintah pusat dan pemerintah daerah. Kabupaten Sleman sendiri telah memiliki sistem Lapor Sleman yang sejalan dengan Smart Regency. Ilmu matematika berperan penting dalam pemecahan masalah terkait dengan pemodelan topik yang sering dilaporkan masyarakat pada masa pasca pandemi. Oleh karena itu, tujuan penelitian ini untuk mengidentifikasi topik pengaduan masyarakat dengan algoritma *Latent Dirichlet Allocation (LDA)*. LDA adalah suatu metode untuk pemodelan topik berdasarkan konsep probabilistik yang mencari kemiripan dalam sebuah dokumen yang kemudian dikelompokkan menjadi beberapa topik. Data yang digunakan pada penelitian ini merupakan data sekunder dari Kominfo Sleman. Tahapan penelitian diawali dengan pengambilan data, dilanjutkan dengan *preprocessing* data. Terdapat beberapa tahap *preprocessing* diantaranya *case folding*, *remove punctuation*, *stopword removal*, *tokenizing*, dan *stemming*. Tahap terakhir adalah *topic modelling* dengan algoritma LDA. Hasil analisis menunjukkan bahwa pemodelan LDA dengan 5 topik yang memiliki nilai *coherence* terbesar. Hasil pemodelan memberikan gambaran bahwa topik yang paling sering muncul dalam laporan pengaduan masyarakat terhadap pemerintah Sleman adalah pelayanan, jalan rusak, vaksin, dan terkait informasi yang dibutuhkan masyarakat.

**Kata kunci:** *Latent Dirichlet Allocation (LDA)*, *matematika*, *pengaduan masyarakat*.

## I. PENDAHULUAN

Partisipasi masyarakat memegang peranan penting dalam perkembangan suatu pembangunan. Sebagai penyelenggara pelayanan publik, pemerintah membutuhkan kritik dan saran sebagai tolak ukur keberhasilan pelayanan publik. Penanganan pengaduan masyarakat tersebut akan dimonitor untuk pencatatan, pengarsipan dan pelaporan. Dengan berkembangnya teknologi dan informasi, pemerintah semakin mempermudah penyaluran pendapat masyarakat melalui layanan pengaduan publik yang dapat diakses secara *online*. Salah satu strategi yang telah diaplikasikan di Kabupaten Sleman yaitu Lapor Sleman yang dikelola oleh pihak Dinas Komunikasi dan Informatika Sleman [1].

Tidak dapat dipungkiri bahwa penyelenggaraan pelayanan publik sering mendapatkan kritikan dari masyarakat. Hal ini dilandasi oleh ketidakpuasan masyarakat terhadap pelayanan yang diberikan. Dengan banyaknya aduan dari masyarakat, pemerintah memerlukan pemetaan terkait topik yang *urgent* dan sering menjadi aduan utama dalam pelayanan publik. Dari hasil identifikasi topik tersebut, pemerintah dapat mencari solusi dan bertindak secara cepat dan optimal [2]. Permasalahan umum yang sering dihadapi adalah terbatasnya konversi dari data yang tidak terstruktur menjadi data yang terstruktur. Informasi yang disajikan dalam bentuk teks yang tidak berlabel dan karenanya tidak bisa diatribusikan dalam domain tematik yang sangat banyak. Kategorisasi manual membutuhkan sumber daya manusia dan keuangan yang sangat besar. Oleh karena itu, perlu dilakukan pemodelan topik dalam pengklasifikasian dokumen tersebut [3].

Ilmu statistika memainkan peran penting dalam pengolahan data karena menyediakan berbagai macam pemodelan. Salah satu metode yang dapat menangani data diskrit dan dapat divisualisasi dalam bentuk

grafis adalah metode *Latent Dirichlet Allocation* (LDA) [4]. LDA adalah salah satu algoritma *machine learning* yang tepat untuk melakukan *topic modelling* [5]. *Topic modelling* merupakan sebuah teknik yang bertujuan untuk menemukan pola kata dalam dokumen dengan model probabilitas [6]. Pada penerapan metode LDA setiap kata dalam algoritmanya diasumsikan memiliki tingkat kepentingan yang sama.

Terdapat banyak studi pemodelan topik yang menggunakan algoritma LDA [7]. Beberapa penelitian yang menerapkan metode LDA pada domain data yang berbeda, diantaranya Fernanda (2021) memodelkan persepsi terhadap pembelajaran *online*, Putra (2017) menganalisis topik informasi sosial media yang ada di Surabaya, Rushdi (2022) mengidentifikasi topik berita dalam artikel [8]. Pada penelitian – penelitian di atas metode LDA diterapkan dengan baik untuk analisis pemodelan topik dan similaritas dokumen [9]. Namun, pengaduan masyarakat masih kurang diteliti, sehingga pada penelitian ini bertujuan untuk melakukan pemodelan topik aduan masyarakat dengan metode LDA. Penelitian ini menggunakan representasi grafis hasil LDA dengan pengelompokan topik untuk mengetahui kluster topik yang perlu menjadi perhatian pemerintah untuk meningkatkan pelayanan publik. Hal ini, merupakan *novelty* dari penelitian – penelitian sebelumnya [10].

## II. METODE PENELITIAN

### A. Jenis Penelitian

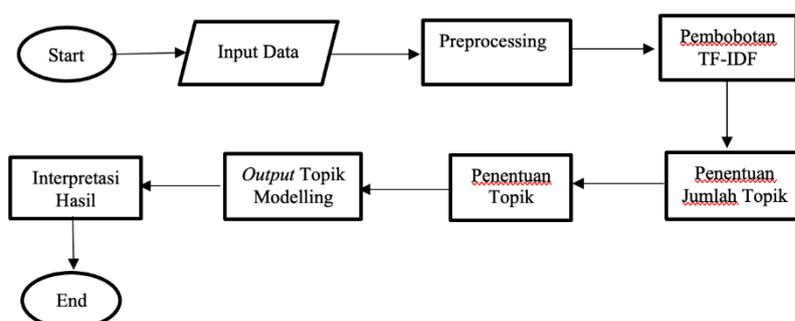
Penelitian ini menggunakan pendekatan kuantitatif. Data yang digunakan merupakan data sekunder yang diperoleh dari Dinas Komunikasi dan Informatika Kabupaten Sleman terkait aduan masyarakat yang disampaikan ke Pemerintah Kabupaten Sleman tahun 2022.

### B. Data Penelitian

Penelitian ini menggunakan data sekunder yang diperoleh dari Dinas Komunikasi dan Informatika Kabupaten Sleman. Variabel yang digunakan adalah topik aduan masyarakat yang disampaikan ke Pemerintah Kabupaten Sleman tahun 2022 dengan total aduan sebanyak 2326 aduan.

### C. Tahapan Penelitian

Tahapan penelitian diawali dengan melakukan *preprocessing* data. Proses *preprocessing* data dilakukan dengan *Python* menggunakan Jupyter – Notebook. Tahap ini diawali dengan menginstall *package* dan mengaktifkan *library* yang akan digunakan. Langkah – langkah *preprocessing* meliputi *case folding*, *remove punctuation*, *stopword removal*, *tokenizing*, dan *stemming* [11]. *Case folding* digunakan untuk mengganti semua huruf dari suatu kata dengan huruf kecil untuk menggeneralisasikan arti kata agar arti kata tidak berubah. Tujuan dari proses *remove punctuation* untuk menghilangkan angka dan tanda baca. Proses *stopword removal* digunakan untuk menghilangkan kata yang tidak diperlukan seperti “di”, “kan”, “oh”, “ke”, “dari”, “mana”. Pada tahapan *tokenizing* berfungsi untuk memisahkan kata dalam kalimat sehingga kata – kata tersebut tidak saling mempengaruhi. Proses *stemming* merupakan proses mengembalikan kata imbuhan pada kata dasarnya [12]. Setelah menyelesaikan tahapan *preprocessing* data, tahapan selanjutnya adalah melakukan pembobotan TF – IDF. Tahapan terakhir yaitu analisis *Topic Modelling* dengan *Latent Dirichlet Allocation* menggunakan *packages Gensim* dan *pyLDAvis* dari *Python*. Seluruh Tahapan penelitian pada penelitian ini disajikan dalam *flowchart* berikut.



GAMBAR 1. DIAGRAM ALIR PENELITIAN

### III. HASIL DAN PEMBAHASAN

#### A. *Cleaning* Data

*Cleaning* merupakan sebuah proses pengolahan dari data mentah menjadi data yang bersih dan dapat digunakan pada pengolahan selanjutnya. Pada proses *cleaning* dilakukan penghapusan tanda baca seperti titik, koma, titik dua, *hashtag*, dan titik koma. Hasil normalisasi data disajikan pada Tabel 1.

TABEL 1. HASIL *CLEANING*

Data Sebelum	Data Sesudah
Pengunjung/pedagang Ps Prambanan Membuang Sampah/benda Dari Lt.4 Ke Pemukiman Warga	Pengunjungpedagang Ps Prambanan Membuang Sampahbenda Dari Lt Ke Pemukiman Warga

Dapat terlihat pada Tabel 1 tanda baca dan angka telah terhapus, sehingga proses selanjutnya akan dilakukan *case folding*.

#### B. *Case Folding*

*Case folding* digunakan untuk mengganti semua huruf dari suatu kata dengan huruf kecil untuk menggeneralisasikan arti kata agar arti kata tidak berubah. Hasil dari proses *case folding* disajikan pada tabel berikut.

TABEL 2. HASIL *CASE FOLDING*

Laporan Pengaduan Sebelum	Laporan Pengaduan Sesudah
Pengunjungpedagang Ps Prambanan Membuang Sampahbenda Dari Lt Ke Pemukiman Warga	pengunjungpedagang ps prambanan membuang sampahbenda dari lt ke pemukiman warga

Pada Tabel 2 di atas telah dilakukan penyamaan huruf menjadi huruf kecil, proses selanjutnya akan dilanjutkan pada *filtering*.

#### C. *Filtering*

*Filtering* adalah sebuah tahapan yang bertujuan untuk mengambil kata yang penting dalam suatu kalimat sehingga mengurangi *noise* (kata yang tidak memiliki makna) dalam satu kalimat. Kata yang tidak ada maknanya tersebut dinamakan *Stopword*. Contoh *Stopword* diantaranya “di”, “kan”, “oh”, “ke”, “dari”, “mana” dan lain – lain. Hasil dari *filtering* disajikan pada tabel berikut.

TABEL 3. HASIL *FILTERING*

Sebelum <i>Filtering</i>	Setelah <i>Filtering</i>
pengunjungpedagang ps prambanan membuang sampahbenda dari lt ke pemukiman warga	pengunjungpedagang ps prambanan membuang sampahbenda lt pemukiman warga

Dari Tabel 3 di atas, telah dilakukan *filtering* pada data, proses selanjutnya akan dilakukan proses *stemming*.

#### D. *Stemming*

Proses *Stemming* merupakan proses mengembalikan kata imbuhan pada kata dasarnya. Hasil dari proses *Stemming* adalah sebagai berikut.

TABEL 4. HASIL *STEMMING*

Sebelum <i>Stemming</i>	Setelah <i>Stemming</i>
pengunjungpedagang ps prambanan membuang sampahbenda dari lt ke pemukiman warga	pengunjungpedagang ps prambanan buang sampahbenda lt mukim warga

Setelah dilakukan proses *stemming* pada Tabel 4, selanjutnya akan dilakukan proses *tokenizing*.

E. *Tokenizing*

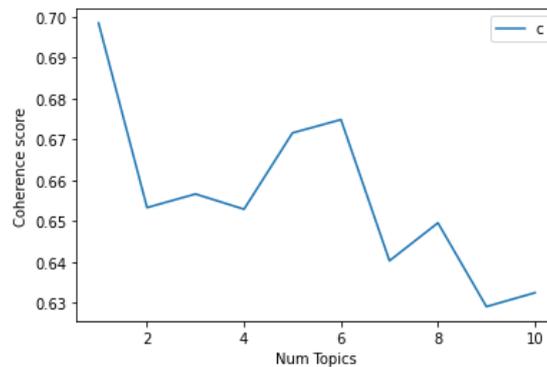
Pada tahapan *tokenizing* berfungsi untuk memisahkan kata dalam kalimat sehingga kata – kata tersebut tidak saling mempengaruhi. Hasil *tokenizing* pada data disajikan pada tabel berikut.

TABEL 5. HASIL *TOKENIZING*

Sebelum <i>Tokenizing</i>	Setelah <i>Tokenizing</i>
pengunjungpedagang ps prambanan membuang sampahbenda dari lt ke pemukiman warga	'pengunjungpedagang' 'ps' 'prambanan', 'buang' 'sampahbenda' 'lt' 'mukim' 'warga'

F. *Topic Modelling*

Hasil *topic modelling* divisualisasikan dengan grafis *coherence* dari data pengaduan masyarakat Kabupaten Sleman tahun 2022 untuk menentukan jumlah topik yang dihasilkan. Dengan menggunakan LDA, peneliti menggunakan nilai topik sejumlah 10 seperti gambar berikut.



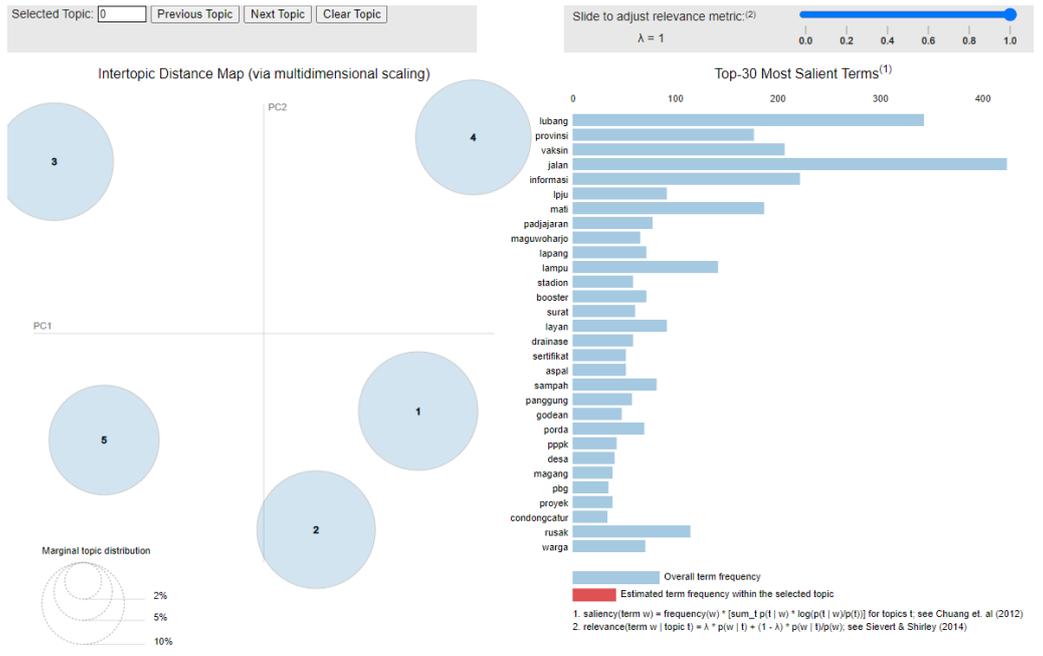
GAMBAR 2. GRAFIK *COHERENCE SCORE*

Pada Gambar 2 di atas diperoleh bahwa nilai *coherence score* terbaik terletak pada jumlah topik 1 dengan nilai *coherence score* tertinggi pada jumlah topik ke 1 adalah 0,698513. Dikarenakan jumlah topik 1 terlalu luas, Selanjutnya, dilakukan pemilihan nilai *coherence score* untuk melakukan *topic modelling*. Nilai *num topic value* dapat dilihat pada tabel berikut.

TABEL 6. HASIL *COHERENCE VALUE* TOPIK PENGADUAN

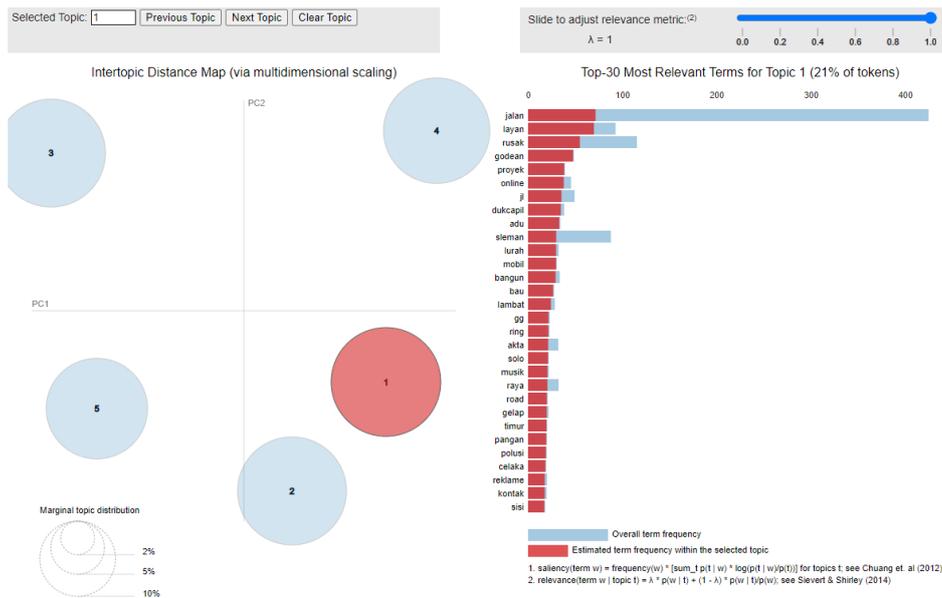
Num Value	Coherence Value
1	0,698513
2	0,653289
3	0,656613
4	0,652884
5	0,671593
6	0,674849
7	0,640249
8	0,649537
9	0,629015
10	0,632407

*Num topic* dengan jumlah 10 memiliki peluang tertinggi untuk muncul dalam dokumen. Selanjutnya, dari 10 topik tersebut, dipilih *coherence score* pada *num value* 5 sehingga didapatkan 5 *term*.



GAMBAR 3. GRAFIK NUM TOPIC

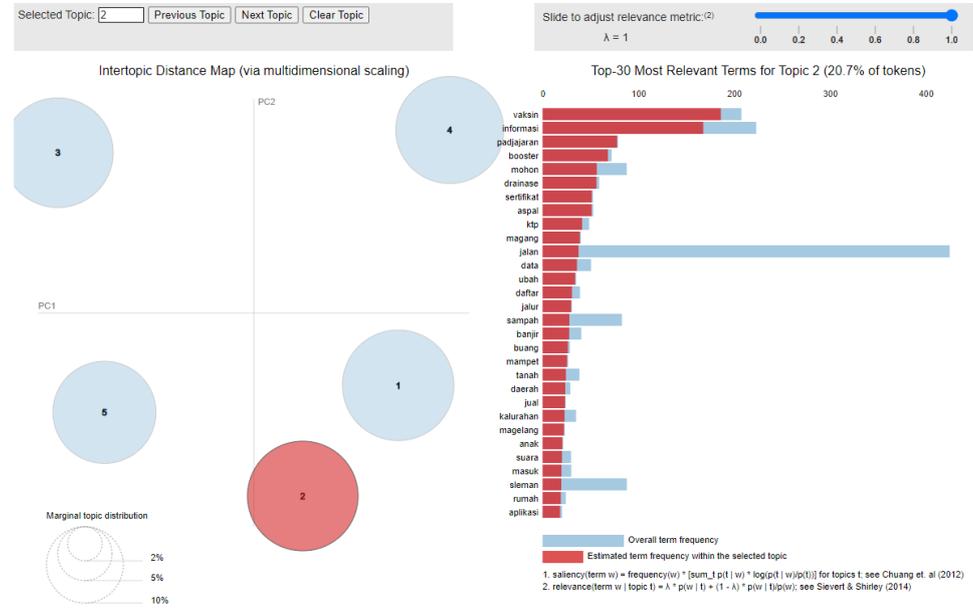
Pada Gambar 3, ukuran gelembung pada setiap topik memiliki kesamaan disebabkan oleh bobot kata pada setiap topik hampir sama sehingga tidak ada perbedaan pada ukuran gelembung. Letak koordinat pada setiap gelembung ditentukan oleh nilai bobot yang telah dihitung. Pada bagian sebelah kiri, terdapat *bar plot* yang memvisualisasikan kata yang ditampilkan beserta frekuensi yang muncul pada topik tersebut dan berjumlah 30 kata yang *urgent* pada seluruh topik.



GAMBAR 4. GRAFIK NUM TOPIC 1

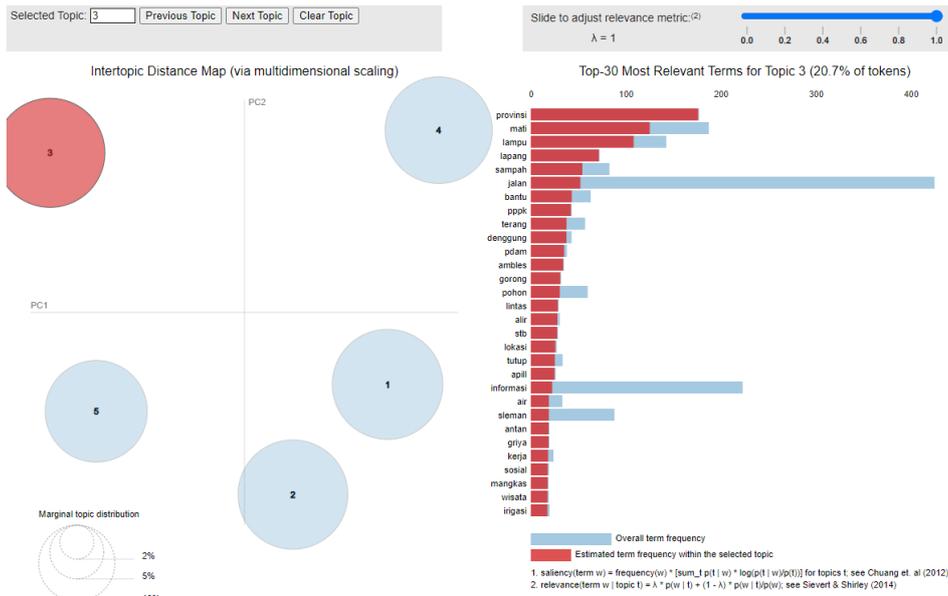
Gambar 4 merupakan visualisasi pada topik pertama. Grafik berwarna merah memvisualisasikan jumlah kemunculan kata topik 1 dan grafik warna biru memvisualisasi kemunculan kata pada keseluruhan dokumen. Kata yang muncul pada topik pertama diurutkan berdasarkan pada frekuensi terbanyak. Kata

tersebut adalah jalan, layan, rusak, godean, proyek, online, jl, dukcapil, adu, sleman, lurah, mobil, bangun, bau, lambat, gg, ring, akta, solo, music, raya, road, gelap, timur, pangan, polusi, celaka, reklame, kontak, sisi. Pada grafik disisi kanan terdapat angka 21% tokens yang mengartikan dari keseluruhan dokumen mewakili topik 1 sebesar 21% yang relevan. Pada topik pertama mewakili bahwa aduan yang masuk adalah terkait jalan yang rusak dengan pembangunan yang lambat.



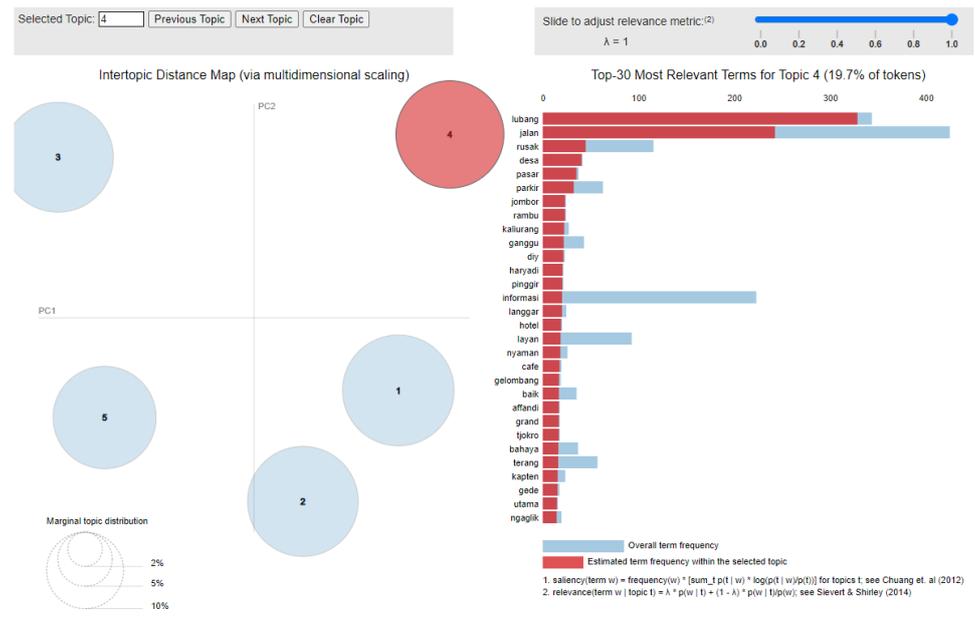
GAMBAR 5. GRAFIK NUM TOPIC 2

Kata yang muncul pada topik kedua yang diurutkan berdasarkan frekuensi terbanyak yaitu vaksin, informasi, padjajaran, booster, mohon, drainase, sertifikat, aspal, ktp, magang, jalan, data, ubah, daftar, jalur, sampah, banjir, buang, mampet, tanah, daerah, jual, kalurahan, magelang, anak, suara, masuk, sleman, rumah, aplikasi. sisi. Pada grafik disisi kanan terdapat angka 20,7% tokens yang mengartikan dari keseluruhan dokumen mewakili topik 2 sebesar 20,7% yang relevan. Topik kedua merepresentasikan tentang masalah yaitu vaksin booster. Selain itu, terdapat kata buang sampah dan banjir yang berarti terdapat masalah lingkungan yang harus segera diselesaikan.



GAMBAR 6. GRAFIK NUM TOPIC 3

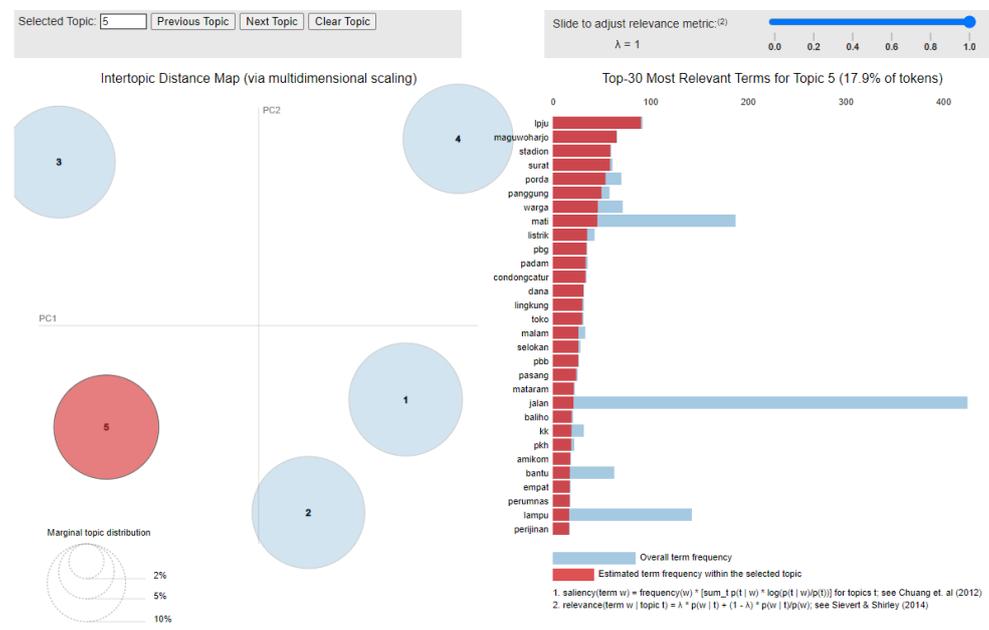
Kata yang muncul pada topik ketiga yang diurutkan berdasarkan frekuensi terbanyak yaitu provinsi, mati, lampu, lapang, sampah, jalan, bantu, pppk, terang, dengging, pdam, ambles, gorong, pohon, lintas, air, stb, lokasi, tutup, apil, informasi, air, Sleman, antan, griya, kerja, sosial, mangkas, wisata, irigasi. Pada grafik disisi kanan terdapat angka 20,7% tokens yang mengartikan dari keseluruhan dokumen mewakili topik 3 sebesar 20,7% yang relevan. Hasil dari topik ketiga mempertegas topik kedua. Pada topik ketiga terdapat kata pdam air dan gorong ambles. Kata tersebut merepresentasikan bahwa masyarakat mengalami kendala terkait air pdam dan gorong – gorong yang ambles.



GAMBAR 7. GRAFIK NUM TOPIC 4

Kata yang muncul pada topik keempat yang diurutkan berdasarkan frekuensi terbanyak yaitu lubang, jalan, rusak, desa, pasar, parkir, jombor, rambu, kaliurang, ganggu, diy, haryadi, pinggir, informasi, langgar,

hotel, layan, nyaman, cafe, gelombang, baik, affandi, grand, tjokro, bahaya, terang, kapten, gede, utama, ngaglik. Pada grafik disisi kanan terdapat angka 19,7% tokens yang mengartikan dari keseluruhan dokumen mewakili topik 3 sebesar 19,7% yang relevan. Hasil dari topik keempat merepresentasikan mengenai fasilitas umum seperti jalan lubang dan rusak, parkir, dan lampu lalu lintas yang masih sangat perlu dilakukan perbaikan.



GAMBAR 8. GRAFIK NUM TOPIC 5

Kata yang muncul pada topik kelima yang diurutkan berdasarkan frekuensi terbanyak yaitu lpu, maguwoharjo, stadion, surat, porda, panggung, warga, mati, listrik, pbg, padam, condongcat, dana, lingkung, toko, malam, selokan, pbb, pasang, mataram, jalan, baliho, kk, pkh, amikom, bantu, empat, perumnas, lampu, perijinan. Pada grafik disisi kanan terdapat angka 17,9% tokens yang mengartikan dari keseluruhan dokumen mewakili topik 5 sebesar 17,9% yang relevan. Topik kelima mewakili aduan masyarakat terkait lpu, mati listrik yang masih menjadi kendala utama dalam pelayanan publik.

Hasil topik pengaduan masyarakat dengan *Latent Dirichlet Allocation* menggunakan nilai *perplexity* dengan jumlah topik lima dan jumlah minimal kata adalah sepuluh. Proses perhitungan *perplexity* menggunakan *package* gensim dengan algoritma *python* secara otomatis. Nilai *perplexity* berguna untuk menentukan topik dengan menggunakan nilai yang terbesar. Hasil perhitungannya adalah sebagai berikut.

Topic: 0 Word:  $0.028 * \text{"jalan"} + 0.027 * \text{"layan"} + 0.021 * \text{"rusak"} + 0.019 * \text{"godean"} + 0.015 * \text{"proyek"} + 0.015 * \text{"online"} + 0.014 * \text{"jl"} + 0.014 * \text{"dukcapil"} + 0.013 * \text{"adu"} + 0.012 * \text{"sleman"}$

Topic: 1 Word:  $0.074 * \text{"vaksin"} + 0.066 * \text{"informasi"} + 0.031 * \text{"padjajaran"} + 0.027 * \text{"booster"} + 0.022 * \text{"mohon"} + 0.022 * \text{"drainase"} + 0.020 * \text{"sertifikat"} + 0.020 * \text{"aspal"} + 0.016 * \text{"ktp"} + 0.015 * \text{"magang"}$

Topic: 2 Word:  $0.041 * \text{"lpu"} + 0.030 * \text{"maguwoharjo"} + 0.027 * \text{"stadion"} + 0.027 * \text{"surat"} + 0.025 * \text{"porda"} + 0.023 * \text{"panggung"} + 0.021 * \text{"warga"} + 0.021 * \text{"mati"} + 0.016 * \text{"listrik"} + 0.016 * \text{"pbg"}$

Topic: 3 Word:  $0.136 * \text{"lubang"} + 0.101 * \text{"jalan"} + 0.019 * \text{"rusak"} + 0.017 * \text{"desa"} + 0.015 * \text{"pasar"} + 0.013 * \text{"parkir"} + 0.010 * \text{"jombor"} + 0.010 * \text{"rambu"} + 0.009 * \text{"kaliurang"} + 0.009 * \text{"gangu"}$

Topic: 4 Word:  $0.070 * \text{"provinsi"} + 0.050 * \text{"mati"} + 0.043 * \text{"lampu"} + 0.028 * \text{"lapang"} + 0.022 * \text{"sampah"} + 0.021 * \text{"jalan"} + 0.017 * \text{"bantu"} + 0.017 * \text{"pppk"} + 0.015 * \text{"terang"} + 0.015 * \text{"dengung"}$

Hasil identifikasi topik divisualisasikan menggunakan *Wordcloud* berdasarkan pada frekuensi kemunculan kata dalam *corpus*. Hal tersebut dapat memudahkan analisis topik dengan cara melihat kata yang paling banyak muncul pada setiap topik. Dari hasil perhitungan menggunakan metode LDA diperoleh lima gambaran yang direpresentasikan pada gambar berikut.



GAMBAR 9. STOPWORD 1



GAMBAR 10. STOPWORD 2



GAMBAR 11. STOPWORD 3



GAMBAR 12. STOPWORD 4



GAMBAR 13. STOPWORD 5

Gambar *stopword* 1 kata yang dominan muncul adalah jalan, rusak, layan, dukcapil, proyek, godean dan *online*. Gambar *stopword* 2 kata yang dominan muncul adalah informasi, vaksin, sertifikat, drainase, booster, dan padjajaran. Gambar *stopword* 3 kata yang dominan muncul adalah maguwoharjo, stadion, porda, warga, listrik, dan padam. Gambar *stopword* 4 kata yang dominan muncul adalah jalan, lubang, pasar, terang, dan informasi. Gambar *stopword* 5 kata yang dominan muncul adalah provinsi, mati, lampu, lapang, dan sampah.

#### IV. SIMPULAN DAN SARAN

Dari hasil penelitian didapatkan pemodelan topik dengan metode LDA pada dataset laporan pengaduan masyarakat tahun 2022 disimpulkan bahwa pengaduan masyarakat lebih banyak tentang pelayanan, jalan rusak, vaksin, dan terkait informasi yang dibutuhkan masyarakat. Dari pemetaan topik tersebut, dapat digunakan oleh pemerintah Kabupaten Sleman dalam melakukan evaluasi dan perbaikan dengan memaksimalkan pelayanan serta mengatasi masalah yang *urgent* dan harus segera diselesaikan. Hasil dari probabilitas kemunculan pada setiap kata dipengaruhi oleh banyaknya topik dan dokumen. Dari penelitian di atas, didapatkan 5 model topik yang divisualisasikan dengan bentuk *bar chart* untuk mengetahui topik utama terkait masalah pengaduan masyarakat di Kabupaten Sleman tahun 2022. Saran yang dapat diberikan yaitu penelitian ini dapat dikembangkan untuk mengetahui tingkat akurasi dengan mempertimbangkan nilai  $\alpha$  dan  $\beta$  pada pencarian topik dengan LDA.

#### DAFTAR PUSTAKA

- [1] Destarani.A.R, Slamet. I, and Subanti.S, “Trend Topic Analysis Using Latent Dirichlet Allocation (LDA) Study Case : Denpasar People’s Complaints Online Website,” vol. 5. Jurnal Ilmiah Teknik Elektro Komputer dan Informatika (JITEKI), 2019, pp. 50 – 58.
- [2] ER.N.A.S, “Implementasi Latent Dirichlet Allocation (LDA) Untuk Klasterisasi Cerita Berbahasa Bali,”vol. 8. Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK), 2019, pp. 127 – 134.
- [3] Fernanda. J.W, “Pemodelan Persepsi Pembelajaran Online Menggunakan Latent Dirichlet Allocation,” vol. 9. Statistika, 2021, pp. 79 – 85.
- [4] Karmila.A, and Sapruni, “Pengelolaan Pengaduan Masyarakat dalam Pelayanan Publik,” vol. 18. Jurnal Ilmiah Administratie, 2022, pp. 25 – 31.
- [5] Karmila.S and Ardianti.V.I, “Metode Latent Dirichlrt Allocation Untuk Menentukan Topik Teks Suatu Berita”, vol. 16. Jurnal Informatika & Komputasi, 2022, pp. 36 – 44.
- [6] Putra. I.M.K.B and Kusumawardani, “Analisis Topik Informasi Publik Media Sosial di Surabaya Menggunakan Pemodelan Latent Dirichlet Allocation (LDA),” vol. 6. Jurnal Teknik ITS, 2017, pp. A311 – A316.
- [7] Rushdi.V.F, and Sari.I, “Identifikasi Topik Artikel Berita Menggunakan Topic Modelling Dengan Latent Dirichlet Allocation”, vol. 27. Jurnal Ilmiah Informatika Komputer, 2022, pp. 169 – 176.

- [8] Santoso.Y.O, and Nugroho.R.S.A, “Pengelompokan Jurnal Ilmiah Berdasarkan Judul Menggunakan LDA,” vol. 3. Jurnal Proxies, 2019, pp. 32 – 42.
- [9] Setijohatmo.U.T, Rachmat.S, Susilawati.T, and Rahman Y, “Analisis Metoda Latent Dirichlet Allocation Untuk Klasifikasi Dokumen Laporan Tugas Akhir Berdasarkan Pemodelan Topik,” vol. 11. Industrial Reseach Workshop and National Seminar, 2020.
- [10] Siringoringo.R, Jamaluddin, and Perangin – Angin.R, “Pemodelan Topik Berita Mnggunakan Latent Dirichlet Allocation dan K – Means Clustering,”vol. 4. Jurnal Informatika Kaputama, 2020, pp. 216 – 222.
- [11] Suparyati, and Utami.E, “Pengamatan Tren Ulasan Hotel Menggunakan Pemodelan Topik Berbasis Latent Dirichlet Allocation,” vol. 6. Jurnal Informatika dan Komputer, 2022, pp. 169 – 175.
- [12] Yohanitas.W.A, “Strategi Penanganan Pengaduan dalam Rangka Peningkatan Pelayanan Publik,” vol. 15. Jurnal Ilmu Administasi (JIA), 2018, pp. 103 – 115.