

PEMODELEN TOPIK PENELITIAN BIDANG KEPERAWATAN INDONESIA PADA *REPOSITORY* JURNAL SINTA MENGGUNAKAN METODE *TOPIC MODELLING LDA (LATENT DIRICHLET ALLOCATION)*

Yoga Sahria¹, Nurul Isnaini Febriarini², Pamulatsih Dwi Oktavianti³

¹Program Studi Informatika Fakultas Teknologi Informasi dan Teknologi
Universitas Teknologi Yogyakarta

^{2,3}Program Studi Sarjana Keperawatan (STIKES) Al Islam Yogyakarta

¹Jl. Siliwangi Jl. Ring Road Utara, Jombor Lor, Sendangadi, Kec. Mlati, Kabupaten Sleman,
Daerah Istimewa Yogyakarta 55285

^{2,3}Jl. Bantul, Dukuh MJ.I/1221 Yogyakarta. Telp. (0274) 373454, 381489

e-mail: ¹yogasahria11@gmail.com, ²nurulif@gmail.com ,

³pamulatsihdwioktavianti@gmail.com

Abstrak

Pada saat ini penelitian di bidang kesehatan semakin meningkat baik dari peneliti, pemerintah, akademis bahkan dari kalangan umum. Penelitian dosen, mahasiswa, peneliti dan lain-lain sudah terdigitalisasi dalam repository Jurnal. Adapun Penelitian kesehatan khususnya keperawatan semakin hari semakin bertambah dan menyebabkan kesulitan dalam mencari topik penelitian yang sesuai dengan kebutuhan, sehingga dibutuhkan metode yang dapat membantu untuk mencari topik penelitian yang mengorganisasi informasi penelitian keperawatan yang digunakan untuk referensi pencarian terkait topik penelitian keperawatan secara umum. Researcher dan akademis menerbitkan artikel penelitian merupakan salah satu program tri dharma perguruan tinggi sehingga setiap hari, bulan dan tahunnya bertambah artikel penelitian di dalam repository jurnal. Untuk mencari dan menemukan penelitian sekarang sangat mudah dapat ditemukan di repository OJS (Open Jurnal System), tetapi permasalahannya adalah mengetahui bagaimana tren di bidang keperawatan berdasarkan tahun. Metode yang digunakan dalam penelitian ini yaitu metode topic modelling LDA (Latent Dirichlet Allocation). Penelitian ini dilakukan dengan tujuan untuk mengidentifikasi bagaimana metode topic modelling LDA dapat melakukan analisis tren topik penelitian dengan pemodelan topik terhadap judul-judul penelitian di bidang keperawatan di Indonesia yang diperoleh dari Repository Jurnal SINTA. Penelitian ini menjadi referensi dalam melakukan penelitian keperawatan di Indonesia berdasarkan topik yang sudah dimodelkan.

Kata Kunci: *topic modelling, keperawatan, LDA*

Abstract

At this time research in the health sector is increasing both from researchers, government, academics and even from the general public. Research lecturers, students, researchers and others have been digitalized. As for health research, especially nursing, is increasing day by day and causes difficulty in finding research topics that suit your needs, so a method is needed that can help to find research topics that organize nursing research information that is used for search references related to nursing research topics in general. Researchers and academics publish research articles as one of the Tri Dharma College programs so that every day, month and year there are more research articles in the journal repository. It is very easy to find and find research now in the OJS repository (Open Journal System), but the problem is knowing how the trends in nursing are based on the

year. The method used in this research is the LDA (Latent Dirichlet Allocation) topic modeling method. This research was conducted with the aim of identifying how the LDA topic modeling method can analyze research topic trends by modeling the topics of research titles in the field of nursing in Indonesia obtained from the SINTA Journal Repository. This research become a reference in conducting nursing research in Indonesia based on the topics that have been modeled.

Keywords: topic modelling, Nursing, LDA

1. PENDAHULUAN

Keperawatan merupakan sebuah profesi yang unik memiliki fokus utama caring, yaitu bagaimana proses untuk memberikan dan mengelola asuhan yang dibutuhkan seorang pasien [1]. Hal ini menjadikan seorang profesi perawat memiliki peran baik untuk negara dan berperasan sebagai pemberi asuhan sebagai kemampuan klinis dan juga koordinator sebagai komponen manajerial untuk merawat pasien. Tingkat pendidikan keperawatan di Indonesia telah mengalami perkembangan yang cukup pesat, mulai dari awalnya hanya berfungsi sebagai pembantu tenaga medis hingga sekarang sebagai profesi yang professional dalam memberikan perawatan kepada pasien. Penelitian di bidang keperawatanpun semakin pesatnya perkembangan zaman semakin hari semakin bertambah banyak dalam mengembangkan sebuah keilmuan keperawatan di Indonesia[2].

Penelitian di Bidang keperawatan menjadi entitas yang penting dalam dunia akademik. Kemajuan penelitian dapat dilihat dari segi jumlah penelitian dan arah topik penelitian yang dibahas[3]. Penelitian di Bidang Keperawatan setia hari, bulan dan tahunnya pasti akan bertambah artikel penelitian di dalam repository jurnal. Untuk mencari dan menemukan penelitian bagi akademis, dosen, mahasiswa, pemerintah bahkan orang umum sekarang sangat mudah dapat ditemukan pada repository OJS (*Open Jurnal System*) di SINTA. Namun untuk melihat topik arah penelitian di Bidang keperawatan yang terdapat di *repository* perlu membaca satu persatu publikasi ilmiah dari seluruh author (penulis), hal tersebut tentu akan memerlukan usaha, waktu dan tenaga. Oleh karena itu diperlukan suatu metode yang bertujuan untuk mengotomisasi pembacaan arah tren topik penelitian di Bidang keperawatan di suatu *repository* SINTA.

Langkah otomatisasi yang dapat dilakukan untuk mencari sebuah topik penelitian menggunakan pendekatan topic modelling. *Topic Modelling* merupakan suatu pendekatan untuk menganalisis sebuah kumpulan berbentuk dokumen teks dan dapat mengelompokkan menjadi beberapa topik [4]. Metode yang digunakan pada penelitian ini yaitu dengan *topic modelling Latent Dirichlet Allocation* (LDA). Pemodelan topik LDA adalah metode Model statistik dari sekumpulan ide untuk mempresentasikan dan menentukan topik menjadi distribusi kosa kata[5]. Tujuan dari topic modelling LDA yaitu menentukan topik secara otomatis dari sekumpulan dokumen[6]. LDA merupakan model probabilistik generatif dari koleksi data diskrit seperti korpus teks. Dasar pemikiran atau ide mengapa ingin melakukan penelitian ini untuk mengetahui tren topik penelitian di Bidang Keperawatan di Indonesia.

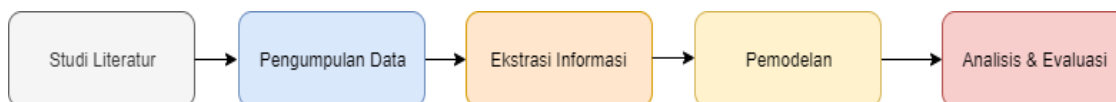
Metode penelitian yang digunakan untuk melakukan *topic modeling* yang paling populer adalah *Latent Dirichlet Allocation* (LDA). Metode LDA telah banyak dikembangkan oleh *author* dalam menganalisis topik dari sebuah teks maupun dokumen. Penelitian sebelumnya terkait dengan menggunakan metode LDA yakni penelitian menerapkan Metode *Bayesian* dan Model *Latent Dirichlet Allocation* (LDA) Di Media Sosial, data yang digunakan diambil dari twitter kompas.com hasil dari penelitian[7] dapat melihat tren berita masyarakat terkait dengan penyebaran gambar atau video lokasi bom di Surabaya

dengan nilai probabilitas topik 0.10057. Selanjutnya penelitian Pemodelan Topik Pengguna di Twitter Mengenai Aplikasi “Ruangguru”, data diambil dari twitter komentar mengenai pendapat pengguna tentang aplikasi ruang guru[8] hasil dari penelitian tersebut dapat mengelompokkan 28 topik yang sering diperbincangkan dalam aplikasi ruang guru. Penelitian selanjutnya menganalisis klasifikasi opini tweet pada media sosial twitter menggunakan *Latent Dirichlet Allocation* (LDA) data diperoleh dari twitter [9], Hasil penelitian tersebut nilai akurasi 66% dengan nilai sensitivitas sebesar 77%. Penelitian selanjutnya aplikasi latent dirichlet allocation (lda) pada clustering data teks [10] adapun data yang diperoleh dari twitter dengan keyword “Bandung”, hasil dari penelitian ini menghasilkan 24 buah topik. Penelitian terkait selanjutnya topic modelling skripsi menggunakan metode *latent dirichlet allocation* [11].

Berdasarkan hasil analisis beberapa penelitian pada metode topic modelling LDA, penulis pada penelitian ini juga menerapkan metode LDA akan tetapi domainnya berbeda yaitu di Bidang Keperawatan. Dalam pemodelan topik penelitian ini berdasarkan data judul-judul publikasi yang diambil dari kumpulan repositori jurnal di SINTA khusus di Bidang Keperawatan. Setelah data didapatkan selanjutnya diproses menggunakan metode LDA, hasilnya menunjukkan apa saja topik-topik penelitian yang dibahas pada penelitian tersebut, kata-kata apa saja yang sering muncul dan digunakan pada penelitian tersebut, dan apa saja topik yang sering dibahas oleh *author* di Bidang Keperawatan Indonesia.

2. METODE PENELITIAN

Pada langkah – langkah penelitian ini menguraikan tahapan-tahapan yang dilakukan pada proses penelitian agar sesuai dan berjalan dengan baik sehingga akan mencapai tujuan yang diinginkan. Metodologi yang digunakan dalam penelitian sesuai dengan judul adalah “*Pemodelan Topik Penelitian Bidang Keperawatan Indonesia Pada Repository Jurnal SINTA Menggunakan Metode Topic Modelling LDA (Latent Dirichlet Allocation)*”. Adapun metodologi dan tahap-tahap yang dilakukan dalam penelitian dapat dilihat pada Gambar 1 sebagai berikut:



Gambar 1. Metode Penelitian

2.1. Studi Literatur

Pada Studi literatur dalam penelitian ini yaitu cara yang dipakai untuk menghimpun data-data atau sumber-sumber yang berhubungan dengan topik dalam suatu penelitian. Studi literatur yang didapatkan dari berbagai sumber, jurnal, buku dokumentasi, internet dan pustaka.

2.2. Pengumpulan Data

Pada tahap ini dilakukan pengumpulan data tentang topik penelitian terkait dengan keperawatan. Semua tahap pada proses pengumpulan data tersebut diperoleh dari Website Jurnal yang berkaitan tentang Penelitian Keperawatan di Indonesia. Website jurnal tentang kesehatan tersebut dilakukan *scraping* untuk mengambil data dari jurnal tersebut berdasarkan *keyword* topik dan afiliasi di Jurnal SINTA.

2.3. Ekstrasi Informasi

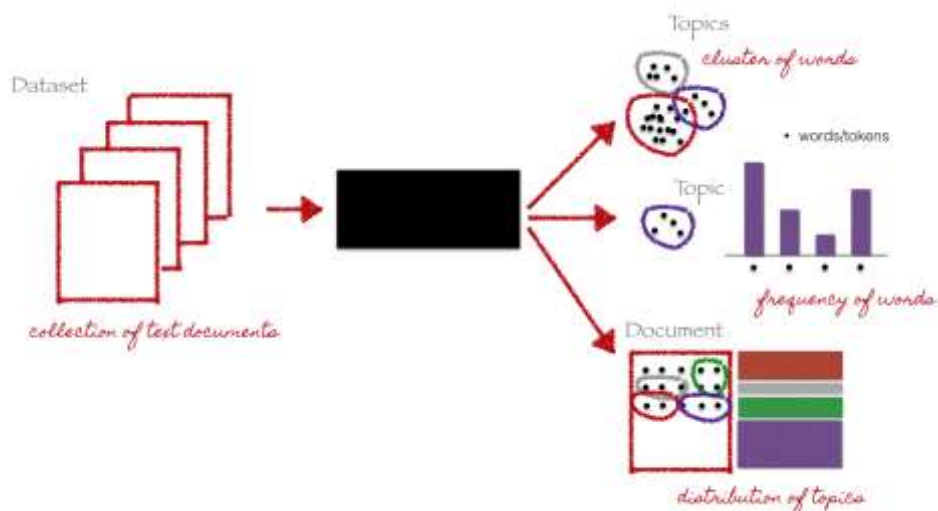
Ekstraksi informasi dalam penelitian ini yaitu untuk pengolahan bahasa alami yang belum terstruktur, dengan cara mengubah teks tidak terstruktur menjadi informasi dalam bentuk terstruktur. Berbagai jenis informasi di Internet ditransmisikan secara tidak terstruktur melalui website jurnal domain di Indonesia, menyebabkan munculnya

kebutuhan akan suatu teknologi untuk menganalisa teks dan menemukan pengetahuan yang relevan dalam bentuk informasi terstruktur dengan menggunakan metode LDA.

2.4. Pemodelan

Pemodelan topik adalah metode untuk menemukan topik dan tren dari *corpus*. Pemodelan topik telah diterapkan di beberapa bidang. Metode pemodelan topik umumnya digunakan untuk secara otomatis mengatur, memahami, mencari, dan meringkas arsip elektronik yang besar. "Topik" menandakan hubungan variabel yang tersembunyi, untuk diperkirakan, yang menghubungkan kata-kata dalam kosa kata dan kemunculannya dalam dokumen. Sebuah dokumen terdiri macam-macam topik campuran. Pemodelan topik ini dapat menemukan tema yang tersembunyi melalui koleksi dokumen dan dapat mengetahui keterangan dokumen sesuai dengan tema tersebut. Setiap kata dilihat dan sebagian diambil dari salah satu topik tersebut. Akhirnya, sebuah dokumen cakupan dari distribusi topik dihasilkan dan menyediakan cara baru untuk mengeksplorasi data tentang perspektif topik.

Pemodelan pada penelitian ini menggunakan metode LDA (*Latent Dirichlet Allocation*). LDA adalah Pemodelan topik probabilistik yang setiap dokumen direpresentasikan secara acak atau campuran lebih dari satu set topik laten dan masing-masing topik direpresentasikan sebagai distribusi atas kosa kata. Dasar ide pemodelan *topic modeling* yaitu sebuah topik yang terdiri dari kata-kata tertentu yang dapat menyusun topik tersebut dari dokumen-dokumen. Pemodelan topik pada penelitian ini yaitu digunakan untuk menemukan topik dalam penelitian keperawatan di Indonesia berdasarkan judul jurnal. Pemodelan topik dapat menggambarkan makna dari dokumen secara semantik yang tersembunyi dalam teks yang jumlahnya besar dan dapat menemukan informasi dari data teks yang tidak terstruktur. Menurut blei Gambar 2 di bawah ini menjelaskan cara kerja LDA.



Gambar 2. Cara kerja LDA

Pada Gambar 2 menurut blei LDA mengasumsikan proses generatif berikut adalah rumus matematis untuk setiap dokumen w dalam sebuah *corpus* D adalah sbb:

1. Pilih $N \sim \text{Poisson}(\xi)$,
2. Pilih $\theta \sim \text{Dir}(\alpha)$,
3. Untuk setiap N kata w_n ,

- a. Pilih Topik $z_n \sim \text{Multinomial}(\theta)$,
- b. Pilih sebuah kata w_n dari $p(w_n | z_n, \beta)$.

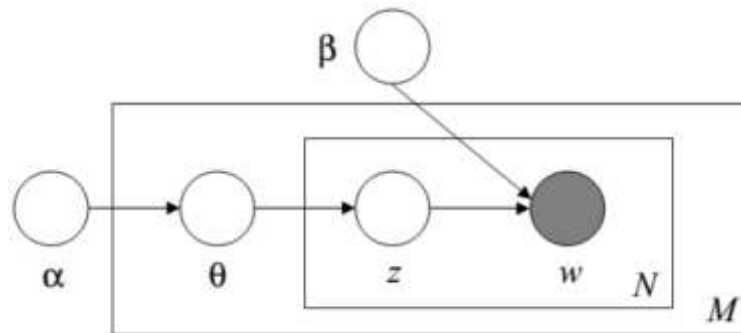
Beberapa asumsi penyederhanaan yang dibuat didalam distribusi dari (latent) topik bahwa diketahui mengikuti k distribusi Dirichlet. Kedua, probabilitas kata adalah matriks β berukuran $k \times V$ yang mana $\beta_{ij} = p(w^j=1|z^i=1)$. Sedangkan k sebagai distribusi Dirichlet memiliki fungsi densitas dapat dilihat pada persamaan (1) sebagai berikut:

$$a. \quad p(\theta|\alpha) = \frac{\Gamma(\sum_{i=1}^k \alpha_i)}{\prod_{i=1}^k \Gamma(\alpha_i)} \theta_1^{\alpha_1-1} \dots \theta_k^{\alpha_k-1} \quad (1)$$

Adapun bentuk dalam distribusi bersama dari Topik *mixture* θ dari N topik Z dan N kata w besyarat α dan β dapat dilihat pada persamaan (2) sebagai berikut:

$$b. \quad p(\theta, \mathbf{z}, \mathbf{w}|\alpha, \beta) = p(\theta|\alpha) \prod_{n=1}^N p(z_n|\theta) p(w_n|z_n, \beta) \quad (2)$$

Bentuk dari representasi model LDA dapat digambarkan dalam sebuah diagram dapat dilihat pada Gambar 3.



Gambar 3. Representasi Model LDA

Bentuk distribusi marginal dari $p(\mathbf{w} | \alpha, \beta)$ didapat dengan menintegrasikan persamaan (2) terhadap θ dapat menghasilkan persamaan (3):

$$p(\mathbf{w}|\alpha, \beta) = \int p(\theta|\alpha) \left(\prod_{n=1}^N \sum_{z_n} p(z_n|\theta) p(w_n|z_n, \beta) \right) d\theta \quad (3)$$

Akhirnya, diperoleh perkalian densitas marginal untuk sebuah dokumen yang akan memperoleh probabilitas marginal sebuah *corpus* persamaan (4) sebagai berikut:

$$p(\mathbf{D}|\alpha, \beta) = \prod_{d=1}^M \int p(\theta_d|\alpha) \left(\prod_{n=1}^{N_d} \sum_{z_n} p(z_{dn}|\theta_{dn}, \beta) \right) d\theta_d \quad (4)$$

Keterangan:

- ξ : adalah *Poisson* kategorikal distribusi memakai variabel random diskrit
- Dir* : adalah *Dirichlet*
- θ : adalah distribusi topik dokumen
- w^j : adalah identitas semua kata dalam semua dokumen j
- z^i : adalah topik untuk kata dalam dokumen i
- z^{ij} : adalah topik untuk kata ke-j dalam dokumen i
- w^{ij} : adalah kata spesifik
- θ_i : adalah distribusi topik untuk dokumen i

- θ_d : adalah level variabel dokumen dan sampel diambil sekali per-dokumen
- φ_k : adalah distribusi kata untuk topik k
- α_i : adalah parameter dari *Dirichlet* sebelum distribusi topik per-dokumen i
- α_k : adalah parameter dari *Dirichlet* jumlah topik per-dokumen k
- w_n : adalah identitas semua kata dalam semua dokumen n
- z_n : adalah identitas topik dari semua kata dalam semua dokumen n

- β_{ij} : adalah parameter dari *Dirichlet* ke- j dalam parameter per topik i
 w_{dn} : adalah *word-level* variabel dan sampel diambil sekali kata disetiap *doc*
 z_{dn} : adalah *topic-level* variabel dan sampel diambil sekali kata disetiap *doc*
 M : adalah menunjukkan jumlah dokumen
 N : adalah jumlah kata dalam dokumen yang diberikan
 K : adalah jumlah topik
 W : adalah identitas semua kata dalam semua dokumen
 V : adalah jumlah kata dalam kosakata
 Z : adalah identitas topik dari semua kata dalam semua dokumen
 α : adalah parameter dari *Dirichlet* sebelum distribusi topik per-dokumen α_i
 β : adalah parameter dari *Dirichlet* sebelum distribusi kata per topik

2.5. Analisis dan Evaluasi

Dalam proses analisis data penelitian ini adanya diskusi menurut sumber dari kemenkes, WHO dan sumber lainnya dengan memperhatikan hasil penelitian yang dibuat. Evaluasi hasil penelitian yang dihasilkan dengan menggunakan metode evaluasi SUS (*System Usability Scale*). SUS atau *System Usability Scale* merupakan kuesioner untuk mengukur persepsi tingkat kegunaan. *System Usability Scale* (SUS) berisi 10 pertanyaan dimana partisipan diberikan pilihan skala 1–5 untuk dijawab berdasarkan pada seberapa banyak mereka setuju dengan setiap pernyataan tersebut terhadap produk atau fitur yang kita uji. Nilai 1 berarti sangat tidak setuju dan 5 berarti sangat setuju dengan pernyataan tersebut.

3. HASIL DAN PEMBAHASAN

Bagian hasil dan Pembahasan memuat hasil-hasil dari penelitian serta pembahasan menyeluruh dari masing-masing hasil yang didapatkan dari penelitian yang dibahas. Dalam melakukan analisis *topic modelling* diperlukan tahapan-tahapan secara runtut agar hasil pemodelan yang dihasilkan baik. Berikut tahap-tahap implementasi yang dilakukan sebagai berikut:

3.1. Pengumpulan Data

Tahap pengumpulan data dalam penelitian ini yaitu diambil judul penelitian khusus dibidang keperawatan di *repository* jurnal SINTA. Data yang diambil yaitu dengan mengakses <http://sinta.ristekbrin.go.id/> kemudian mengklik *source journal* pada menu jurnal SINTA, kemudian mencari nama jurnal dengan menggunakan *keyword* “keperawatan” pada kolom pencarian jurnal. Hasil dari pencarian Nama Jurnal Keperawatan di Indonesia terdapat sejumlah 50 jurnal dari berbagai afiliasi. Adapun cuplikan hasil *scraping* nama jurnal dapat dilihat pada Table 1.

Tabel 1. Cuplikan Tabel Hasil Scraping

Nama Jurnal	Afiliasi	ISSBN
Jurnal Keperawatan Indonesia	Universitas Indonesia	2354-9203
Jurnal Keperawatan Padjadjaran	Universitas Padjadjaran	2442-7276
NERS:Jurnal Keperawatan	Universitas Andalas	2461-0747
Jurnal Smart Keperawatan	(STIKes) Karya Husada	2502-5236
Media Keperawatan Indonesia	UNIMUS	2615-1669
Jurnal Ilmiah Keperawatan Imelda	Univ Imelda Medan	2597-7172
Jurnal Keperawatan Terapan (e-Journal)	Politeknik Kesehatan Kemenkes	2442-6873
Jurnal Ilmu Keperawatan Jiwa	(PPNI) Jawa Tengah	2621-2978
Jurnal Ilmu Keperawatan Anak	Universitas Andalas	2461-0747
Jurnal Ilmu Keperawatan dan Kebidanan	STIKES Muhammadiyah Kudus	2442-9902
Jurnal Keperawatan	STIKES Kendal	6212-96X
Dst	dst	dst


```

str = re.sub(r'[\w]|\_',' ',str)
#remove digit from string
str = re.sub("\S*\d\S*", "", str).strip()
#remove digit or numbers
str = re.sub(r"\b\d+\b", " ", str)
#to lowercase
str = str.lower()
#Remove additional white spaces
str = re.sub('[\s]+', ' ', str)
return str
def preprocessing(str):
str = removeSentence(str)
str = cleaning(str)
str = removeStopword(str)
return str

```

Adapun hasilnya dari script tersebut dapat dilihat pada tabel 3 sebagai berikut:

Tabel 3. Judul Penelitian setelah di Preprocessing Data

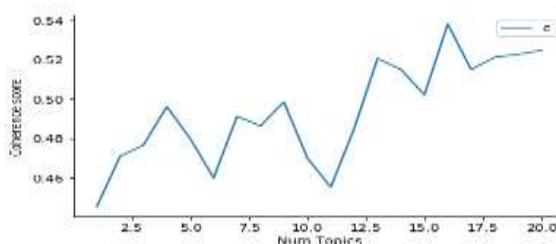
No	Judul Penelitian
1	Faktor-faktor yang mempengaruhi kejadian TB paru dan upaya penanggulangannya
2	Analisis kualitatif bakteri koliform pada depo air minum isi ulang di kota Singaraja Bali
3	Prevalensi Cacing Pada Murid Sekolah Dasar Wajib Belajar Pelayanan Gerakan Terpadu Pengentasan Kemiskinan Daerah Kumuh Di Wilayah DKI Jakarta
4	Pengamatan tempat perindukan Aedes aegypti pada tempat penampungan air rumah tangga pada masyarakat pengguna air olahan
5	Analisis kualitatif bakteri koliform pada depo air minum isi ulang di kota singaraja bali
6	Potensi daun pandan wangi untuk membunuh larva nyamuk Aedes aegypti
7	Hubungan Faktor Lingkungan Rumah dengan Penularan TB Paru Kontak Serumah
8	Habitat Perkembangbiakan Dan Aktivitas Menggigit Nyamuk Anopheles Sundaicus Dan Anopheles Subpictus Di Purworejo, Jawa Tengah
9	Studi Komunitas Nyamuk Di Desa Sebusus (Daerah Endemis Filariasis), Sumatera Selatan Tahun 2004 (A study on mosquito community)
10	Analisis dan manajemen risiko kesehatan pencemaran udara studi kasus di sembilan kota besar padat transportasi

Dst dst

3.4. Pemodelan Topik LDA

Dalam penelitian ini pada tahap pemodelan topik LDA memanfaatkan *library* yang disediakan oleh *python* yaitu dengan menginstal *pip install LDA*. langkah pertama yaitu menampilkan nilai kohorensi setiap kata pada judul penelitian keperawatan di Indonesia. Kemudian untuk melihat persebaran kata pada setiap topik menggunakan algoritma TF-IDF (*Term Frequency – Inverse Document Frequency*) untuk pembobotan kata. Untuk Mencari topik yang dominan atau banyak muncul dengan algoritma TF-IDF dengan probabilitas kemunculan topik katanya yang dominan dalam judul penelitian keperawatan di Indonesia adapun hasilnya dapat dilihat pada Tabel 3. Berikut adalah hasil grafik nilai koherenasi setiap kata yang dihasilkan dapat dilihat pada Gambar 3.

Adapun hasil dari visualisasi grafik nilai koherenasi setiap kata dapat dilihat pada gambar 5 sebagai berikut:



Gambar5. Grafik nilai koherenasi setiap kata

Tabel 4. Hasil Nilai Kohorensi Pada Setiap Topik

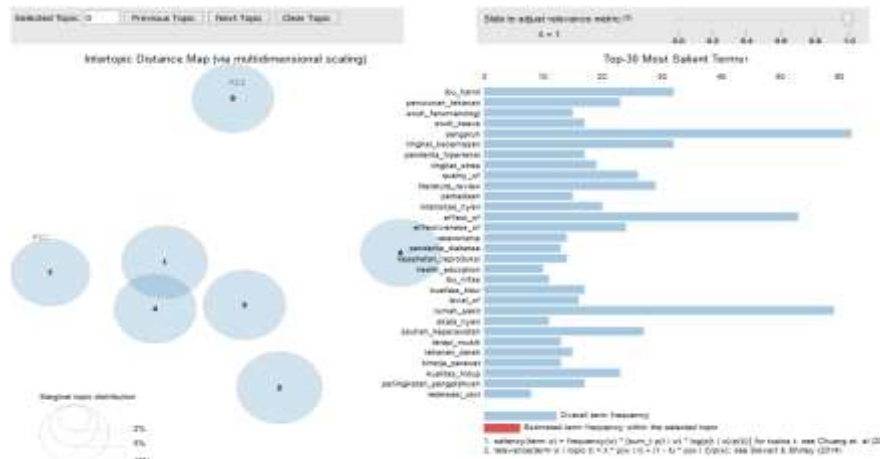
Topik	Kata
Topik 1	Word: 0.011*"pengaruh" + 0.010*"effect_of" + 0.009*"of" + 0.009*"self_efficacy" + 0.009*"kualitas_tidur" + 0.008*"penderita_diabetes" + 0.008*"relationship" + 0.007*"tekanan_darah" + 0.006*"diabetes" + 0.006*"wilayah_kerja"
Topik 2	Word: 0.012*"rumah_sakit" + 0.011*"literature_review" + 0.011*"studi_kasus" + 0.010*"asuhan_keperawatan" + 0.007*"kinerja_perawat" + 0.006*"pengetahuan_sikap" + 0.006*"pengetahuan_ibu" + 0.006*"diabetes_melitus" + 0.006*"systematic_review" + 0.005*"review"
Topik 3	Word: 0.013*"rumah_sakit" + 0.012*"pengaruh" + 0.010*"tingkat_stres" + 0.007*"kualitas_hidup" + 0.007*"skala_nyeri" + 0.006*"rawat_inap" + 0.006*"faktor_faktor" + 0.005*"lanjut_usia" + 0.005*"kepuasan_pasien" + 0.005*"sectio_caesarea"
Topik 4	Word: 0.014*"of" + 0.009*"quality_of" + 0.009*"intensitas_nyeri" + 0.008*"wilayah_kerja" + 0.008*"peningkatan_pengetahuan" + 0.007*"ibu_nifas" + 0.007*"health" + 0.006*"rumah_sakit" + 0.006*"pengalaman" + 0.006*"kerja_perawat"
Topik 5	Word: 0.012*"effectiveness_of" + 0.012*"effect_of" + 0.010*"of" + 0.009*"perbedaan" + 0.008*"level_of" + 0.007*"health_education" + 0.007*"terapi_musik" + 0.006*"gangguan_jiwa" + 0.006*"mahasiswa_keperawatan" + 0.006*"rumah_sakit"
Topik 6	Word: 0.016*"ibu_hamil" + 0.010*"of" + 0.009*"effect_of" + 0.008*"anak_usia" + 0.008*"kesehatan_reproduksi" + 0.008*"pasien_diabetes" + 0.008*"patients_with" + 0.007*"pasien_gagal" + 0.007*"pendidikan_kesehatan" + 0.006*"pasien"
Topik 7	Word: 0.015*"pengaruh" + 0.014*"penurunan_tekanan" + 0.013*"tingkat_kecemasan" + 0.012*"studi_fenomenologi" + 0.011*"penderita_hipertensi" + 0.008*"hipertensi" + 0.008*"of" + 0.007*"self_care" + 0.006*"lansia" + 0.006*"tingkat_pengetahuan"

3.5 Visualisasi Pemodelan Topik LDA

Visualisasi pemodelan topik pada penelitian ini setelah selesai melakukan pemodelan topik LDA, model TF-IDF disimpan ke dalam bentuk *pyLDavis* yang dapat membentuk visualisasi dari masing-masing topik dan kata yang paling banyak muncul dalam judul penelitian kedalam diagram adapun bentuk visualisasi dapat dilihat pada gambar 5.

Pada gambar 6 menggambarkan 30 kata penting yang muncul di *corpus*. Panel kanan menggambarkan tentang kata-kata dominan yang dibahas topik dari dataset judul penelitian keperawatan di Indonesia. Berdasarkan panel kanan visualisasi ibu hamil, fenomenologi, tingkat kecemasan, penderita hipertensi, tingkat stress, intensitas nyeri, penderita diabetes, kesehatan reproduksi, ibu nifas, kualitas tidur, asuhan keperawatan, tekanan darah, kinerja perawat, kualitas hidup, relaksasi otot menjadi kata yang banyak muncul di *corpus*. Oleh karena itu, peneliti menyimpulkan bahwa judul-judul penelitian keperawatan di Indonesia berdasarkan visualisasi dengan menggunakan library *python*

pyLdavis topik penelitian keperawatan di Indonesia yang terindeks SINTA yaitu berkaitan dengan 30 kata penting yang muncul pada visualisasi pada gambar 5.



Gambar 6. Visualisasi Topik Dengan *pyLDavis*

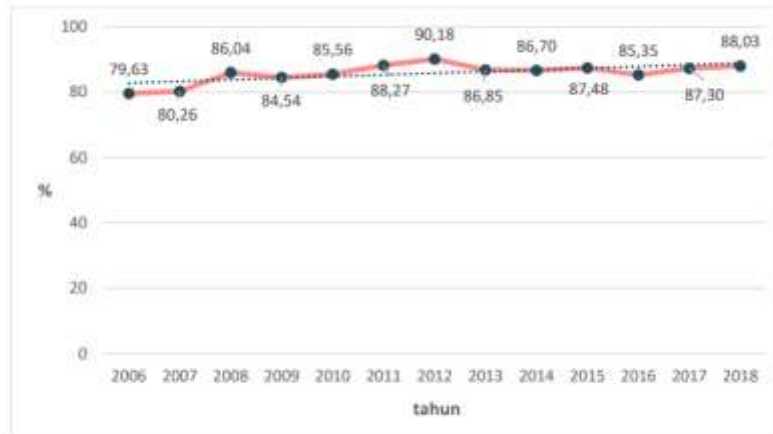
Visualisasi selanjutnya dengan menggunakan visualisasi *wordcloud*. *Word cloud* (atau sering disebut juga sebagai *tag cloud* yaitu dapat merepresentasi visual dari data teks. *Word cloud* dimanfaatkan untuk menggambarkan metadata kumpulan kata kunci atau *keywords* atau *tags* pada sebuah dokumen untuk memberikan visualisasi topik bentuk teks. Visualisasi *wordcloud* semakin besar katanya semakin banyak muncul kata teksnya. Visualisasi *wordcloud* penelitian ini berdasarkan hasil pemodelan LDA pada tabel 3 sehingga terdapat 7 Visualisasi *wordcloud* yang menunjukkan hasil pemodelan topik. Visualisasi *wordcloud* dianggap cukup representatif untuk menampilkan hasil pemodelan topik karena dapat mengetahui kemunculan *term-term* yang sering muncul berdasarkan topik yang dimodelkan [13]. Adapun visualisasi *word cloud* berdasarkan topik yang terbentuk dapat dilihat pada gambar 7.



Gambar 7. Visualisasi Topik Dengan Word cloud

3.6 Analisis dan Evaluasi

Berdasarkan implementasi *topic modelling* LDA berikut hasil pemodelan topik yaitu diantaranya adalah hamil, fenomenologi, tingkat kecemasan, penderita hipertensi, tingkat stress, intensitas nyeri, penderita diabetes, kesehatan reproduksi, ibu nifas, kualitas tidur, asuhan keperawatan, tekanan darah, kinerja perawat, kualitas hidup, relaksasi otot. Adapun topik yang paling banyak diteliti adalah Ibu Hamil. Hal ini terbukti menurut kemenkes (kementerian kesehatan) Indonesia Program terkait ibu hamil sangat diperhatikan dalam pelayanan kesehatan Ibu dan anak. Adapun Cakupan Pelayanan Kesehatan Ibu Hamil K4 Di Indonesia dapat dilihat pada Gambar 8.

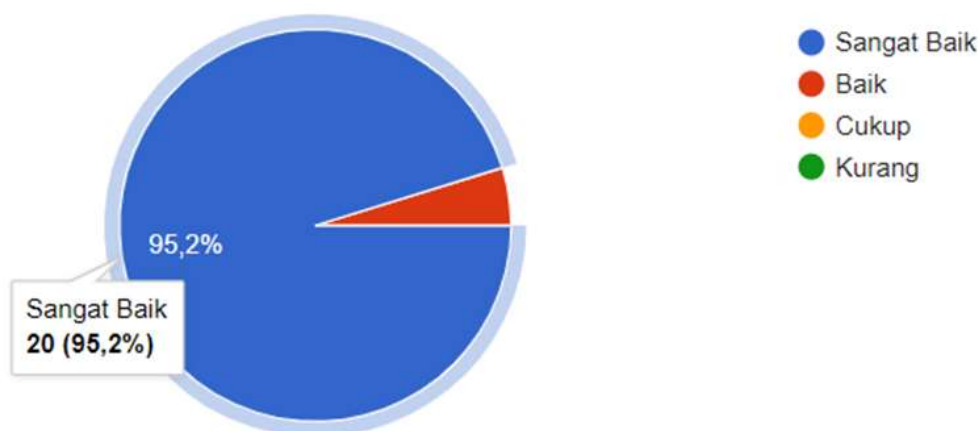


Sumber: Ditjen Kesehatan Masyarakat, Kemenkes RI, 2019

Gambar 8. Cakupan Pelayanan Kesehatan Ibu Hamil K4 Di Indonesia

Selama tahun 2006 sampai tahun 2018 cakupan pelayanan kesehatan ibu hamil K4 cenderung meningkat. Jika dibandingkan dengan target Rencana Strategis (Renstra) Kementerian Kesehatan tahun 2018 yang sebesar 78%, capaian tahun 2018 telah mencapai target yaitu sebesar 88,03%.

Adapun hasil pengujian penelitian ini menggunakan SUS (*System Usability Scale*) yang disampaikan responden peneliti, tenaga kesehatan, dan akademis rata-rata menjawab dengan pemodelan topik yang dihasilkan yaitu 95,2% mengatakan sangat baik dan 4,8 % baik adapun prosentasi disajikan pada Gambar 9 sebagai berikut:



Gambar 9. Hasil Pengujian Menggunakan SUS (*System Usability Scale*)

4. KESIMPULAN

Penelitian ini berhasil memodelkan topik penelitian keperawatan di Indonesia yang diperoleh dari repository SINTA. Penelitian ini juga menampilkan visualisasi hasil pemodelan topik kedalam visualisasi *wordcloud* dan *pyLDAvis* terhadap hasil topik persebaran kata terhadap judul penelitian keperawatan di Indonesia diantaranya yaitu tentang ibu hamil, fenomenologi, tingkat kecemasan, penderita hipertensi, tingkat stress, intensitas nyeri, penderita diabetes, kesehatan reproduksi, ibu nifas, kualitas tidur, asuhan keperawatan, tekanan darah, kinerja perawat, kualitas hidup, relaksasi otot dan lain-lain. Saran dari penelitian ini selanjutnya dapat dibuat aplikasi berbasis android atau website dengan fasilitas *dashboard* untuk memodelkan topik secara *real time* yang bertujuan untuk melihat topik penelitian-penelitian kesehatan di Indonesia khususnya dibidang keperawatan sehingga berguna bagi peneliti, pemerintah sebagai pengambilan keputusan terkait rencana kerja pemerintah.

UCAPAN TERIMAKASIH

Terimakasih Kepada LPMI STIKES Al Islam Yogyakarta dan LPPM Universitas Teknologi Yogyakarta yang telah mendukung pendanaan penelitian ini sehingga dapat diselesaikan dengan baik. Terimakasih kepada dosen keperawatan STIKES AL Islam Yogyakarta yang membantu diskusi dalam menyelesaikan penelitian ini.

DAFTAR PUSTAKA

- [1] D. Koerniawan, N. Daeli, and . S., "Aplikasi Standar Proses Keperawatan: Diagnosis, Outcome, Dan Intervensi Pada Asuhan Keperawatan," *J. Keperawatan Silampari*, vol. 3, no. 2, pp. 739-751, 2020.
- [2] C. Casman, A. Ahadi Pradana, E. Edianto, and L. O. Abdul Rahman, "Kaleidoskop Menuju Seperempat Abad Pendidikan Keperawatan di Indonesia," *J. Endur.*, vol. 5, no. 1, p. 115, 2020, doi: 10.22216/jen.v5i1.4291.
- [3] A. Nurlyayli and M. A. Nasichuddin, "Topic Modeling Penelitian Dosen JPTEI UNY pada Google Scholar Menggunakan Latent Dirichlet Allocation," vol. 4, no. November, pp. 154-161, 2019, doi: 10.21831/elinvo.v4i2.
- [4] F. Martin and M. Johnson, "More Efficient Topic Modelling Through a Noun Only Approach," *Proc. Australas. Lang. Technol. Assoc. Work.*, pp. 111-115, 2015, [Online]. Available: <http://www.nltk.org/howto/wordnet.html>.
- [5] A. F. Hidayatullah and M. R. Ma'Arif, "Road traffic topic modeling on Twitter using latent dirichlet allocation," *Proc. - 2017 Int. Conf. Sustain. Inf. Eng. Technol. SIET 2017*, vol. 2018-Janua, no. November 2017, pp. 47-52, 2018, doi: 10.1109/SIET.2017.8304107.
- [6] R. Kusumaningrum, M. I. A. Wiedjayanto, S. Adhy, and Suryono, "Classification of Indonesian news articles based on Latent Dirichlet Allocation," *Proc. 2016 Int. Conf. Data Softw. Eng. ICoDSE 2016*, pp. 1-5, 2017, doi: 10.1109/ICODSE.2016.7936106.
- [7] M. Fajriyanto, "Penerapan metode bayesian dalam model latent dirichlet allocation di media sosial application of bayesian methods in latent dirichlet allocation model in social media," pp. 1-6, 2018.
- [8] B. W. Arianto and G. Anuraga, "Topic Modeling for Twitter Users Regarding the 'Ruangguru' Application," *J. ILMU DASAR*, vol. 21, no. 2, p. 149, 2020.
- [9] M. Fitriasih and R. Kusumaningrum, "Analisis Klasifikasi Opini Tweet Pada Media Sosial Twitter Menggunakan Latent Dirichlet Allocation (LDA)," *Semin. Nas. Teknol. Inf. dan Komun. 2019 (SENTIKA 2019)*, vol. 2019, no. Sentika, pp. 177-186, 2019.
- [10] Zulhanif, Sudartianto, B. Tantular, and I. G. N. M. Jaya, "Aplikasi Latent Dirichlet

- Allocation (Lda) Pada Clustering Data Teks,” *J. Log.*, vol. 7, no. 1, pp. 46–51, 2017.
- [11] A. I. Alfanzar, K. Khalid, and I. S. Rozas, “Topic Modelling Skripsi Menggunakan Metode Latent Dirichlet Allocation,” *JSiI (Jurnal Sist. Informasi)*, vol. 7, no. 1, p. 7, 2020, doi: 10.30656/jsii.v7i1.2036.
- [12] I. Surjandari, A. Rosyidah, Z. Zulkarnain, and E. Laoh, “Mining Web Log Data for News Topic Modeling Using Latent Dirichlet Allocation,” *Proc. - 2018 5th Int. Conf. Inf. Sci. Control Eng. ICISCE 2018*, pp. 331–335, 2019, doi: 10.1109/ICISCE.2018.00076.
- [13] M. F. A. Bashri and R. Kusumaningrum, “Sentiment analysis using Latent Dirichlet Allocation and topic polarity wordcloud visualization,” *2017 5th Int. Conf. Inf. Commun. Technol. ICoICT 2017*, vol. 0, no. c, pp. 4–8, 2017, doi: 10.1109/ICoICT.2017.8074651.



Prosiding- SEMASTER: Seminar Nasional Teknologi Informasi & Ilmu Komputer is licensed under a [Creative Commons Attribution International \(CC BY-SA 4.0\)](https://creativecommons.org/licenses/by-sa/4.0/)