

Topic Modeling in Thesis Titles of Students from the Faculty of Economics Universitas Garut Using Latent Dirichlet Allocation Modeling

Fikri Fahru Roji^{1*}, Dinar Rahayu², Riyad Sabilul Muminin³,
Dendi Ramdani⁴, Dede Hendrik⁵

¹Program Studi Bisnis Digital, Fakultas Ekonomi, Universitas Garut, Garut, Jawa Barat, Indonesia

²Program Studi Sistem Informasi, Institut Pendidikan Indonesia, Garut, Jawa Barat, Indonesia

³⁻⁴Program Studi Rekayasa Perangkat Lunak, Fakultas Komunikasi dan Informasi, Universitas Garut, Garut, Jawa Barat, Indonesia

⁵Program Studi Sistem Informasi, Fakultas Teknik, Institut Teknologi Al-Muhajirin, Purwakarta, Jawa Barat, Indonesia

*Penulis koresponden, e-mail : fikri@uniga.ac.id

Abstract: *In higher education, the completion of a thesis within a 1 -year timeframe is a prerequisite for graduation. The selection of a thesis topic is influenced by personal interest, the expertise of the thesis supervisor, and data availability. This research is designed to analyze the thesis topics of Economics Faculty students at Garut University using the Latent Dirichlet Allocation (LDA) Modeling method. Utilizing quantitative and qualitative approaches, this research applies the concept of big data with techniques such as Data Crawling, Data Preprocessing, and Text Mining. The research successfully conducted topic modeling using the LDA method. The analysis showed that topic modeling with the LDA algorithm resulted in seven common thesis topics used in the students' thesis titles. With this, the research contributes to the understanding and efficacy in the determination of students' thesis topics. It is hoped that the results of this research can be utilized to assist in the efficient completion of theses.*

Keywords: *Topic Analysis; Topic Modeling; Thesis Title; Latent Dirichlet Allocation; LDA*

Abstrak: Dalam pendidikan tinggi, penyelesaian skripsi dalam kurun waktu 1 tahun menjadi prasyarat kelulusan. Pertimbangan penentuan topik skripsi dipengaruhi oleh minat pribadi, keahlian dosen pembimbing, dan ketersediaan data. Penelitian ini dirancang untuk melakukan analisis topik skripsi mahasiswa Fakultas Ekonomi Universitas Garut menggunakan Metode *Latent Dirichlet Allocation (LDA) Modeling*. Menggunakan pendekatan kuantitatif dan kualitatif, penelitian ini mengaplikasikan konsep big data dengan teknik *Crawling Data, Data Preprocessing* dan *Text Mining*. Penelitian ini berhasil dalam melakukan pemodelan topik menggunakan metode LDA. Analisis menunjukkan bahwa pemodelan topik dengan algoritma LDA menghasilkan 7 topik skripsi yang biasa digunakan dalam judul skripsi mahasiswa. Dengan ini, penelitian ini berkontribusi pada pemahaman dan efisiensi dalam penentuan topik skripsi mahasiswa. Diharapkan, hasil penelitian ini dapat dimanfaatkan untuk membantu proses penyelesaian skripsi dalam waktu yang efisien.

Kata kunci: Analisis Topik; Pemodelan Topik; Judul Skripsi; *Latent Dirichlet Allocation*; LDA

PENDAHULUAN

Dalam pertumbuhan dan perkembangan dunia pendidikan, khususnya di tingkat perguruan tinggi, proses penulisan skripsi merupakan suatu unsur yang paling esensial. Skripsi menjadi cerminan penempatan kemampuan dan keilmuan mahasiswa dalam melakukan riset serta

merumuskan dan memahami sebuah topik secara mendalam (Widodo et al., n.d.). Skripsi adalah salah satu metode yang ditentukan sebagai prasyarat kelulusan bagi mahasiswa di berbagai universitas di Indonesia. (Kinasih et al., 2021). Dalam prosesnya, pengerjaan skripsi dilaksanakan maksimal 1 (satu) tahun, jika tidak dapat diselesaikan maka dinyatakan tidak berhak dilanjutkan dan diharuskan mengajukan judul skripsi yang baru (Diah et al., 2020). Batasan waktu maksimal 1 tahun untuk menyelesaikan skripsi merupakan kebijakan yang umum diterapkan di banyak institusi pendidikan tinggi. Hal ini bertujuan untuk memastikan bahwa mahasiswa dapat menyelesaikan tugas akhir mereka dalam waktu yang wajar dan tidak terlalu memanjang.

Salah satu tantangan krusial dalam proses pembuatan skripsi terletak pada penentuan topik skripsi. Beragam faktor dapat mempengaruhi keputusan ini, yang antara lain mencakup minat pribadi mahasiswa, bidang keahlian dosen pembimbing, dan ketersediaan data dan sumber penelitian yang relevan. (Cahyani & Arif, 2022; Safitri et al., 2022; Tafonao et al., 2022; L. A. Utami, 2017; Wangge, 2021). Relevansi dengan bidang studi, ketersediaan data dan sumber penelitian, dan potensi kontribusi terhadap pengetahuan dan praktik di bidang studi menjadi faktor-faktor utama dalam mempertimbangkan topik skripsi (Cahyani & Arif, 2022; Olivia et al., 2022; Wangge, 2021).

Analisis topik skripsi melibatkan proses evaluasi terhadap potensi dan viabilitas dari topik-topik skripsi yang menjadi pilihan mahasiswa. Analisis ini bukan hanya membantu memfasilitasi pemahaman mahasiswa tentang relevansi, keunikan, dan kelayakan topik mereka dalam konteks bidang studi, tetapi juga membantu dalam mempersiapkan mereka secara mental dan emosional dalam menghadapi proses penulisan skripsi (Knafl et al., 2015).

Topic modelling bertujuan untuk menemukan topik dari kumpulan dokumen. *Topic modelling* adalah teknik yang kuat dalam *text mining* yang memungkinkan untuk menemukan pola dan hubungan tersembunyi dalam data dan dokumen teks. Teknik ini telah banyak digunakan dalam berbagai bidang seperti rekayasa perangkat lunak, ilmu politik, ilmu kedokteran, dan linguistik (Jelodar et al., 2018). *Topic modelling* memberikan beberapa manfaat, termasuk mengidentifikasi topik-topik penting, otomatisasi pengkodean dan pengukuran fenomena skala besar, serta membantu dalam interpretasi dan validasi hasil (Masyhuri, 2022).

Dengan mengkaji pemahaman ini, penelitian ini berfokus pada analisis topik skripsi mahasiswa dan memodelkannya menggunakan LDA dengan harapan dapat membantu dalam meningkatkan kualitas penelitian mahasiswa dan efisiensi proses penulisan skripsi.

KAJIAN PUSTAKA

A. Skripsi

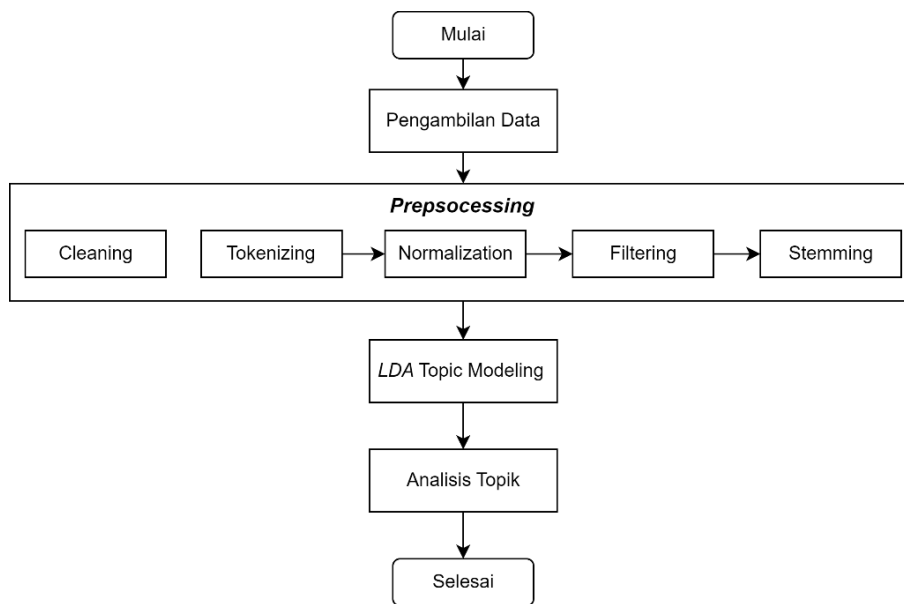
Skripsi merupakan salah satu cara yang ditetapkan sebagai syarat kelulusan bagi mahasiswa di berbagai perguruan tinggi di Indonesia (Kinasih et al., 2021). Dalam prosesnya, pengerjaan skripsi dilaksanakan maksimal 1 (satu) tahun, jika tidak dapat diselesaikan maka dinyatakan tidak berhak dilanjutkan dan diharuskan mengajukan judul skripsi yang baru (LPPM, 2021). Proses bimbingan skripsi merupakan proses kompleks yang melibatkan banyak pihak, termasuk mahasiswa, pembimbing skripsi, dan pemangku kepentingan universitas (Sahratullah et al., 2023; Wangge, 2021; Yarpiransa et al., 2023). Keberhasilan penyelesaian skripsi dalam jangka waktu yang ditentukan membutuhkan koordinasi dan kolaborasi yang efisien di antara pihak-pihak tersebut (Batubara et al., 2022).

B. Pemodelan Topik

Pemodelan topik adalah teknik yang kuat dalam *text mining* yang memungkinkan penemuan pola-pola tersembunyi dan hubungan dalam data dan dokumen teks. Teknik ini telah banyak diterapkan dalam berbagai bidang seperti rekayasa perangkat lunak, ilmu politik, ilmu kedokteran, dan linguistik. Salah satu metode populer untuk pemodelan topik adalah *Latent Dirichlet Allocation* (LDA), yang telah digunakan dan diteliti secara luas dalam bidang ini (Jelodar et al., 2018). Para peneliti telah mengusulkan model-model berbeda berdasarkan LDA untuk meningkatkan pemodelan topik. Pemodelan topik telah diterapkan untuk merekomendasikan artikel ilmiah, mengelompokkan dokumen, meningkatkan pencarian bibliografis, dan menganalisis volume besar informasi yang tak terstruktur (Zuluaga, 2023). Selain itu, pemodelan topik telah digunakan dalam domain khusus seperti literatur hidrologi, ilmu bahan, dan penelitian pemasaran (Masyhuri, 2022). Penerapan pemodelan topik telah menunjukkan hasil yang menjanjikan dalam hal kualitas topik, kemampuan interpretasi, dan performa prediksi (Dieng et al., 2020). Secara keseluruhan, pemodelan topik telah menjadi teknik yang serbaguna dengan berbagai aplikasi dan telah berkontribusi pada kemajuan berbagai bidang penelitian.

METODE PENELITIAN

Sesuai dengan kerangka penelitian pada Gambar 1. Penelitian ini merupakan penelitian kuantitatif dan kualitatif yang menerapkan konsep dari *big data*. *Big data* merujuk pada fenomena data yang melimpah di dunia daring yang ditandai oleh ukurannya yang besar, kecepatannya, dan perbedaannya dari karakteristik basis data konvensional (Zaluchu, 2020). Penggunaan metode campuran, yang melibatkan pengumpulan dan integrasi data kuantitatif dan kualitatif, dapat memberikan hasil yang lebih akurat.



Gambar 1. Metode Penelitian

A. Pengambilan Data

Tahap pertama penelitian adalah pengambilan data dengan mengambil data dari SIMAK UNIGA. Proses pengambilan data menggunakan tool MySQL. Penelitian ini mempunyai batasan masalah yaitu hanya mengambil data judul skripsi yang dilakukan oleh program studi di Fakultas Ekonomi UNIGA.

B. Data Preprocessing

Dalam tahap *Preprocessing*, penelitian ini dijalankan memakai bahasa *Python* dan *Microsoft Excel* sebagai platform penyimpanan data. Proses *preprocessing data* ini penting untuk memfasilitasi *sentiment analysis* dan menaikkan efektivitas sekaligus efisiensi dalam *data processing*. (Nofiyani & Wulandari, 2022). Tahapan proses data *preprocessing* melibatkan beberapa langkah untuk mempersiapkan data sebelum dilakukan Pemodelan Topik. Beberapa langkah yang umum ditempuh dalam pra-pemrosesan data meliputi pelipatan huruf, pemisahan kata, normalisasi, penyaringan (penghapusan kata yang tidak penting), dan pemangkasan kata (N. W. Utami & Artana, 2022). *Transformation Data* dengan tahap *Case Folding*, di mana data-data akan diubah menjadi bentuk yang seragam (huruf kecil). *Data Cleaning* melalui proses *Filtering*, yakni penghapusan kata-kata yang tidak berpengaruh atau tidak informatif. *Tokenizing*, yakni memisahkan setiap kata pada teks dari data yang telah diperoleh. *Normalization* adalah proses mengubah bahasa slang dan kesalahan ketik pada kasus ini, menggunakan Bahasa Indonesia. *Stopwords removal* adalah dengan menghapus kata-kata yang memiliki informasi rendah dalam suatu teks. *Stemming* adalah mencari

bentuk root dari suatu kata. Setelah data dibersihkan, data tersebut disimpan kembali dengan nama databersih dengan ekstensi csv.

C. Pemodelan Topik dengan LDA

Sebelum membuat pemodelan topik dengan LDA, peneliti terlebih dahulu membuat Wordcloud masing masing dari respon positif dan negatif data agar lebih terlihat *corpus* yang ditampilkan dalam data penelitian ini. Pembuatan wordcloud menggunakan software berbasis python. Setelah corpus ditampilkan dalam bentuk Wordcloud, peneliti melakukan pemodelan topik dengan menggunakan LDA. Proses analisis pengelompokan topik dengan LDA menggunakan bantuan software berbasis python. Mengingat *LDA* menerapkan prinsip *partition-based clustering*, peneliti menetapkan terlebih dahulu jumlah *cluster* dari ulasan yang ada. Jumlah *cluster* topik pada penelitian ini mencakup 5 *cluster*, dengan tujuan agar model yang dikembangkan tak menghasilkan topik saling beririsan satu sama lain, atau tidak terlampaui rinci sehingga banyak topik yang harus diinterpretasikan.

D. Analisis Topik

Analisis topik dilakukan berdasarkan pada data keluaran dari tahap sebelumnya. Pada tahap sebelumnya diperoleh grafik data keluaran dari kumpulan penelitian dengan topik tertentu. Analisis topik dilakukan secara subjektif dengan melihat data keluaran. Data keluaran berupa kumpulan kata yang membentuk topik, kemudian setiap dokumen tersebut disesuaikan dengan data keluaran yang memuat dokumen dengan topik. Proses ini menghasilkan deskripsi topik yang bersifat informatif mengenai hal yang dapat mewakili isi dari masing-masing topik tersebut.

HASIL DAN PEMBAHASAN

A. Data Selection

Pengambilan data Skripsi Mahasiswa diambil dari sistem informasi yang berjalan. Data diambil dari basis data institusi dan berfokus pada 2645 judul skripsi. Proses yang diterapkan melibatkan beberapa langkah. Pertama, dilakukan seleksi kriteria yang jelas untuk memilih data yang relevan, seperti tahun kelulusan dan program studi. Kemudian, menggunakan perintah SQL, data-data skripsi dipindahkan dari sistem basis data ke dalam format yang sesuai, yaitu dalam bentuk file CSV. Validasi data dilakukan untuk memastikan integritas dan keakuratan data.

B. Preprocessing

Tahap prapemrosesan, yang merujuk pada persiapan data sebelum analisis menggunakan *Latent Dirichlet Allocation (LDA)*, dilakukan menggunakan aplikasi berbasis web bernama

JupyterLab. Proses ini didahului dengan pengunduhan dan pengaktifan beberapa *library* yang akan digunakan selama tahap ini.

Prosedur prapemrosesan data yang ada terbagi menjadi lima tahapan penting yang mencakup:

1. *Tokenizing* : Tahap ini merujuk pada pemecahan teks ke dalam potongan yang lebih kecil atau "token".
2. *Stopword Removal* : Tahapan yang merujuk pada penghilangan kata-kata yang tidak memiliki informasi penting atau kata-kata yang umum dipakai dan tidak perlu.
3. *Lemmatization* : Proses ini melibatkan pengubahan kata ke bentuk "lemma" atau bentuk dasarnya.
4. *Stemming* : Tahap ini melibatkan pengurangan kata ke bentuk akarnya.

Selama fase ini, informasi dikustomisasi dan diubah agar dapat diproses pada fase selanjutnya. Tabel 1 memberikan ilustrasi hasil dari enam langkah prapemrosesan, yakni:

1. *Case Folding* (Penyatuan Bentuk)
2. *Remove Punctuation* (Menghilangkan Tanda Baca)
3. *Tokenizing* (Pembagian Kalimat jadi Kata)
4. *Normalisasi*
5. *Filtering* (Penyaringan)
6. *Stemming* (Pemotongan Kata ke Bentuk Dasar)

Tabel 1. Hasil Preprocessing

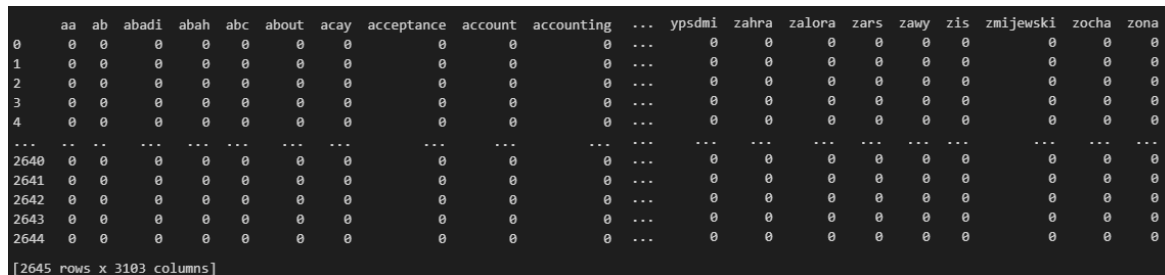
Proses	Hasil
<i>Data Collection</i>	Penerapan Metode Penyusutan Aktiva Tetap Pada PT PLN (Persero) Area Garut
<i>Case Folding</i>	penerapan metode penyusutan aktiva tetap pada pt pln (persero) area garut
<i>Remove Punctuation</i>	penerapan metode penyusutan aktiva tetap pada pt pln persero area garut
<i>Tokenizing</i>	"penerapan" "metode" "penyusutan" "aktiva" "tetap" "pada" "pt" "pln" "persero" "area" "garut"
<i>Normalization</i>	"penerapan" "metode" "penyusutan" "aktiva" "tetap" "pada" "pt" "pln" "persero" "area" "garut"
<i>Filtering</i>	"penerapan" "metode" "penyusutan" "aktiva" "pt" "pln" "persero" "area" "garut"
<i>Stemming</i>	"terap" "metode" "penyusutan" "aktiva" "pt" "pln" "persero" "area" "garut"

C. Topic Modeling dengan LDA

Dari tahap pra-pemrosesan data, kita mampu memperoleh *term* atau matriks yang berisikan kata-kata yang sering muncul. Model *Bag of Words* dipraktikkan untuk mengkalkulasi jumlah kemunculan setiap kata dalam *term* ataupun matriks kata-kata tersebut. Setelah dihitung, hasil frekuensi kemunculan tiap kata digunakan dalam proses distribusi dalam *Latent Dirichlet Allocation* atau *LDA*.

Fase pertama pada proses ini melibatkan penyimpanan kata yang unik yang ada dalam *term* atau matriks kata-kata tersebut. Melalui prosedur ini, ada 3103 kata unik yang berhasil diisolasi.

Langkah kedua dalam proses ini menyangkut kalkulasi kata yang muncul. Dalam menghitung kata tersebut, model *Bag of Words* mengindeksasi tiap kata yang ada sekaligus menghitung frekuensi kemunculannya. *Bag of Words* menghitung berdasarkan kata unik yang telah ditetapkan dalam prosedur sebelumnya. Hasil dari proses *Bag of Words* dapat dilihat dalam Gambar 2.



	aa	ab	abadi	abah	abc	about	acay	acceptance	account	accounting	...	ypsdmi	zahra	zalora	zars	zawy	zis	zmijewski	zocha	zona
0	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0
...																				
2640	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0
2641	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0
2642	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0
2643	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0
2644	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0

[2645 rows x 3103 columns]

Gambar 2. Hasil Proses pada Bag of Words

Setelah melalui serangkaian tahap pra-pemrosesan dan diintegrasikan ke dalam model *Bag of Words*, tahapan selanjutnya ialah melakukan pemodelan topik menggunakan metode *Latent Dirichlet Allocation (LDA)*. Dalam tahap *Bag of Words*, telah muncul sejumlah *token* yang merujuk ke banyaknya kata yang terdapat dalam suatu dokumen. Fungsi *token* ini sangat penting dalam model *LDA* guna melakukan pemodelan.

Dalam proses pemodelan *LDA*, diperlukan serangkaian percobaan untuk dapat menentukan jumlah topik yang paling optimal. Eksperimen tersebut dilakukan dengan cara memvariasikan jumlah topik. Dalam penelitian ini, percobaan dilakukan dengan menggunakan jumlah topik yang beragam, yaitu: 2, 3, 4, 5, 6, 7, 8, dan 9.

D. Analisis Topik

Selanjutnya, analisis dilakukan terhadap keseluruhan model *LDA* berdasarkan berbagai eksperimen dengan jumlah topik yang berbeda-beda. Hasil dari serangkaian percobaan tersebut menunjukkan bahwa jumlah topik yang optimal adalah 8 topik. Ouput dari pemodelan topik dengan menggunakan *LDA* dapat dibaca pada Tabel 2. Semakin tinggi nilai koherensi dari sebuah topik, maka semakin mudah untuk memahami maknanya berdasarkan kata-kata yang membentuk topik tersebut. Dengan kata lain, semakin sering kata-kata dalam topik tersebut muncul secara simultan, maka nilai koherensi topik tersebut akan semakin tinggi. Berdasarkan Tabel 3, kita dapat menginterpretasikan 8 topik dari model yang telah dibuat. Hasil interpretasi topik ini digunakan untuk melihat pola dan tren dari ulasan aplikasi SatuSehat oleh pengguna.

Tabel 2. Nilai dari Topik Koherensi

Topic	Hasil Nilai-nilai Koherensi	Perkiraan Tema
1	0.053*"pengaruh" + 0.043*"brand" + 0.037*"garut" + 0.025*"bank" + 0.021*"image" + 0.021*"customer" + 0.020*"intention" + 0.013*"syariah" + 0.013*"bpr" + 0.012*"dampak"	Pengaruh Brand Image dalam Industri dan Perbankan di Garut
2	0.056*"garut" + 0.033*"pengaruh" + 0.030*"pajak" + 0.026*"kabupaten" + 0.023*"analisis" + 0.017*"wajib" + 0.015*"studi" + 0.013*"terap" + 0.013*"akuntansi" + 0.011*"patuh"	Analisis Pajak dan Kepatuhan di Kabupaten Garut
3	0.055*"garut" + 0.047*"sistem" + 0.047*"analisis" + 0.035*"kendali" + 0.024*"akuntansi" + 0.022*"pt" + 0.021*"cabang" + 0.019*"kredit" + 0.018*"intern" + 0.018*"anggar"	Sistem Analisis dan Pengendalian di Cabang PT
4	0.137*"kerja" + 0.073*"garut" + 0.070*"pengaruh" + 0.050*"karyawan" + 0.033*"pt" + 0.026*"kabupaten" + 0.023*"motivasi" + 0.021*"lingkung" + 0.020*"pegawai" + 0.019*"disiplin"	Pengaruh Motivasi dan Disiplin Kerja terhadap Karyawan di PT Kabupaten Garut
5	0.046*"pt" + 0.040*"pengaruh" + 0.035*"analisis" + 0.034*"uang" + 0.031*"tbk" + 0.019*"garut" + 0.019*"indonesia" + 0.018*"koperasi" + 0.018*"modal" + 0.018*"kerja"	Pengaruh Keuangan dan Modal terhadap PT, Koperasi, dan Indonesia
6	0.059*"pengaruh" + 0.055*"garut" + 0.051*"beli" + 0.051*"putus" + 0.025*"motor" + 0.022*"of" + 0.019*"wisata" + 0.018*"konsumen" + 0.018*"mouth" + 0.018*"word"	Pengaruh Perilaku Konsumen dalam Keputusan Pembelian dan Promosi Wisata di Garut
7	0.070*"usaha" + 0.037*"daftar" + 0.036*"pengaruh" + 0.030*"indonesia" + 0.025*"efek" + 0.025*"bursa" + 0.019*"nilai" + 0.015*"corporate" + 0.014*"variabel" + 0.012*"social"	Pengaruh Variabel dalam Usaha dan Efeknya terhadap Bursa Efek di Indonesia
8	0.075*"garut" + 0.067*"pengaruh" + 0.064*"kualitas" + 0.050*"faktor" + 0.041*"puas" + 0.041*"layan" + 0.035*"langgan" + 0.035*"produk" + 0.031*"beli" + 0.030*"putus"	Pengaruh Kualitas Produk dan Layanan terhadap Kepuasan Pelanggan di Garut

Topik 1 dengan kata kunci: pengaruh, brand, garut, bank, image, customer, intention, syariah, bpr, dampak. Topik ini berkaitan dengan pengaruh citra merek (brand image) terhadap sektor perbankan di daerah Garut. Pengaruh ini dapat berkaitan dengan persepsi nasabah (customer) dan niat (intention) nasabah untuk berinteraksi dengan bank, terutama dalam konteks perbankan syariah dan BPR.

Topik 2 dengan kata kunci: garut, pengaruh, pajak, kabupaten, analisis, wajib, studi, terap, akuntansi, patuh. Topik ini membahas tentang pengaruh faktor pajak dan analisis terhadap kabupaten Garut. Fokus utamanya adalah pada kepatuhan pajak (tax compliance) di daerah tersebut serta aspek-aspek akuntansi terkait.

Topik 3 dengan kata kunci: garut, sistem, analisis, kendali, akuntansi, pt, cabang, kredit, intern, anggar. Topik ini berkaitan dengan analisis dan pengendalian sistem dalam konteks cabang perusahaan PT di Garut. Hal ini mencakup aspek akuntansi, pengendalian kredit, dan manajemen anggaran.

Topik 4 dengan kata kunci: kerja, garut, pengaruh, karyawan, pt, kabupaten, motivasi, lingkungan, pegawai, disiplin. Topik ini membahas tentang pengaruh motivasi dan disiplin kerja terhadap karyawan di perusahaan PT di Kabupaten Garut. Fokusnya adalah bagaimana faktor-faktor tersebut memengaruhi kinerja karyawan.

Topik 5 dengan kata kunci: pt, pengaruh, analisis, uang, tbk, garut, indonesia, koperasi, modal, kerja. Topik ini berkaitan dengan pengaruh faktor keuangan seperti modal, uang, dan analisis terhadap berbagai entitas, termasuk perusahaan PT, koperasi, dan Indonesia secara keseluruhan.

Topik 6 dengan kata kunci: pengaruh, garut, beli, putus, motor, of, wisata, konsumen, mouth, word. Topik ini membahas pengaruh perilaku konsumen dalam konteks keputusan pembelian, khususnya dalam aspek promosi wisata di Garut. Juga membahas faktor pengaruh "word of mouth" dalam promosi.

Topik 7 dengan kata kunci: usaha, daftar, pengaruh, indonesia, efek, bursa, nilai, corporate, variabel, social. Topik ini membahas pengaruh berbagai variabel dalam konteks usaha, termasuk efeknya terhadap bursa efek di Indonesia. Variabel-variabel tersebut melibatkan nilai perusahaan (corporate values) dan aspek-aspek sosial.

Topik 8 dengan kata kunci: garut, pengaruh, kualitas, faktor, puas, layan, langgan, produk, beli, putus. Topik ini membahas pengaruh kualitas produk dan layanan terhadap kepuasan pelanggan di Garut. Faktor-faktor yang berkontribusi terhadap kepuasan pelanggan, seperti layanan, produk, dan faktor pembelian, juga dianalisis.

KESIMPULAN DAN SARAN

Untuk topik judul skripsi mahasiswa, penelitian ini berhasil melakukan pemodelan topik menggunakan metode *Latent Dirichlet Allocation (LDA)*. Berdasarkan analisis maka dapat disimpulkan bahwa:

1. Pemodelan topik dengan algoritma LDA menghasilkan 7 topik skripsi yang biasa digunakan dalam judul skripsi mahasiswa.
2. Topik yang dihasilkan LDA berkaitan dengan :
 - a. Pengaruh Brand Image dalam Industri dan Perbankan di Garut
 - b. Analisis Pajak dan Kepatuhan di Kabupaten Garut
 - c. Sistem Analisis dan Pengendalian di Cabang PT
 - d. Pengaruh Motivasi dan Disiplin Kerja terhadap Karyawan di PT Kabupaten Garut
 - e. Pengaruh Keuangan dan Modal terhadap PT, Koperasi, dan Indonesia
 - f. Pengaruh Perilaku Konsumen dalam Keputusan Pembelian dan Promosi Wisata di Garut
 - g. Pengaruh Variabel dalam Usaha dan Efeknya terhadap Bursa Efek di Indonesia
 - h. Pengaruh Kualitas Produk dan Layanan terhadap Kepuasan Pelanggan di Garut

DAFTAR PUSTAKA

- Batubara, Megawati, Auliani Meidina, and Shahnaz Salsabila Rahman. 2022. "Menemukan Dukungan Sosial Yang Paling Berkontribusi Terhadap Psychological Well-Being Mahasiswa Yang Sedang Menyusun Skripsi, Dosen Pembimbing Atau Orang Tua?" *Indonesian Psychological Research*.
- Cahyani, Laili, and Muchamad Arif. 2022. "Text Mining Untuk Pengelompokan Skripsi Di Prodi Pendidikan Informatika Universitas Trunojoyo Madura." 8(2): 97–108.

- Diah, Delila Nurbani et al. 2020. “Efek Moderasi Resiliensi Terhadap Hubungan Antara Perfeksionisme Dengan Kecemasan Mengerjakan Skripsi Perfectionism and Anxiety in Working on Mini Thesis.” 6(2): 178–90.
- Dieng, Adji B, Joaquín R Vidal, and David M Blei. 2020. “Topic Modeling in Embedding Spaces.” *Transactions of the Association for Computational Linguistics*.
- Jelodar, Hamed et al. 2018. “Latent Dirichlet Allocation (LDA) and Topic Modeling: Models, Applications, a Survey.” *Multimedia Tools and Applications*.
- Kinasih, Hayu Wikan, Agung Prajanto, and Mila Sartika. 2021. “Peran Dosen Pembimbing Dalam Lulus Tepat Waktu Mahasiswa: Study Pada Mahasiswa Akuntansi Universitas X.” *Proceeding SENDU*: 208–14.
- Knafel, Kathleen et al. 2015. “Delimiting Family in Syntheses of Research on Childhood Chronic Conditions and Family Life.” *Family Process* 54(1): 173–84. <https://onlinelibrary.wiley.com/doi/abs/10.1111/famp.12101>.
- LPPM. 2021. *Pedoman Penulisan Karya Ilmiah Fakultas Ekonomi Universitas Garut*. Garut.
- Masyhuri, Muhammad. 2022. “Topic Modelling Application for Determining Competitiveness Factors of the Small Business Firms.” *International Journal of Social Science and Business*.
- Nofiyani, Nofiyani, and Wulandari Wulandari. 2022. “Implementasi Electronic Data Processing Untuk Meningkatkan Efektifitas Dan Efisiensi Pada Text Mining.” *Jurnal Media Informatika Budidarma*.
- Olivia, Maria et al. 2022. “Relevansi Nilai Informasi Akuntansi Terhadap Harga Saham Pada Beberapa Penelitian Skripsi Dan Jurnal Perguruan Tinggi , Studi Meta Analisis.” 11(September): 1–9.
- Safitri, Alvira Oktavia, Vioreza Dwi Yuniarti, and Deti Rostika. 2022. “Jurnal Basicedu.” 6(4): 7096–7106.
- Sahratullah, Sahratullah et al. 2023. “Analisis Arah Penelitian Skripsi Mahasiswa Program Studi Pendidikan Biologi FKIP Universitas Cordova.” *Jurnal Ilmiah Mandala Education*.
- Tafonao, Talizaro, Ya Gulo, and Tri Murni Situmeang. 2022. “Tantangan Pendidikan Agama Kristen Dalam Menanamkan Nilai-Nilai Kristen Pada Anak Usia Dini Di Era Teknologi.” 6(5): 4847–59.
- Utami, Lilyani Asri. 2017. “Analisis Sentimen Opini Publik Berita Kebakaran Hutan Melalui Komparasi Algoritma Support Vector Machine Dan K-Nearest Neighbor Berbasis Particle Swarm Optimization.” *Jurnal Pilar Nusa Mandiri* 13(1): 103–12. www.tribunnews.com.
- Utami, Nengah W, and Made Artana. 2022. “Text Mining Dalam Analisis Sentimen Pembelajaran Daring Di Masa Pandemi Covid 19 Menggunakan Algoritma K-Nearest Neighbor.” *Jurnal Informatika Teknologi Dan Sains*.
- Wange, Maria Carmelita Tali. 2021. “Penerapan Metode Principal Component Analysis (PCA) Terhadap Faktor-Faktor Yang Mempengaruhi Lamanya Penyelesaian Skripsi Mahasiswa Program Studi Pendidikan Matematika FKIP UNDANA.” *Jurnal Cendekia Jurnal Pendidikan Matematika*.
- Widodo, Arif et al. “Analisis Kemampuan Menulis Makalah Mahasiswa Baru PGSD Universitas Mataram.” VI(1): 77–91.
- Yarpiransa, Yarpiransa, Darjat Saripurna, and Heri Santoso. 2023. “Implementasi Metode Scrum Pada Pengembangan Aplikasi Bimbingan Skripsi Online.” *Hello World Jurnal Ilmu Komputer*.
- Zaluchu, Sonny Eli. 2020. “Strategi Penelitian Kualitatif Dan Kuantitatif Di Dalam Penelitian Agama.” *Evangelikal Jurnal Teologi Injili Dan Pembinaan Warga Jemaat*.
- Zuluaga, Martha. 2023. “Topic Modeling: Perspectives From a Literature Review.” *Ieee Access*.