

Analisis Topik Skripsi Menerapkan Pemodelan *Latent Dirichlet Allocation*

Wahyudi Ariannor^{1*}, Erwin Arry Kusuma², Andita Suci Pratiwi³

Teknik Informatika, STMIK Banjarbaru, Banjarbaru, Indonesia

*e-mail *Corresponding Author*: wahyu.arian@gmail.com

Abstract

At XYZ University, the alignment of thesis titles with the vision of each study program has never been analyzed. The sheer volume of thesis titles makes conventional analysis challenging. Therefore, an automated and rapid analysis technique is needed. Topic modeling with the Latent Dirichlet Allocation (LDA) model offers a solution for facilitating topic analysis. This study compares topic modeling results with and without stemming in data preprocessing. Topic modeling without stemming yielded higher coherence scores compared to stemming, with scores of 0.336 for the Informatics Engineering program and 0.446 for the Information Systems program. Based on the topic modeling analysis, it was found that the thesis topics have a clear focus aligned with the vision and mission of each study program.

Keywords: *Stemming; coherence score; Thesis titles; Vision and mission; Topic modeling*

Abstrak

Pada perguruan tinggi XYZ, judul-judul skripsi yang telah ada belum pernah dilakukan analisis kesesuaian topik dengan visi masing-masing program studi. Jumlah judul-judul skripsi yang begitu banyak akan sulit jika dianalisis secara konvensional. Sehingga diperlukan teknik yang dapat melakukan analisis otomatis dengan cepat. Teknik pemodelan topik dengan model *Latent Dirichlet Allocation* (LDA) dapat menjadi solusi untuk mempermudah menganalisis topik. Penelitian ini membandingkan hasil pemodelan topik antara hasil dengan menerapkan *stemming* dan tanpa *stemming* pada pra-pemrosesan data. Pemodelan topik tanpa proses *stemming* menghasilkan *Coherence Score* yang lebih tinggi dibandingkan dengan proses *stemming*, yaitu 0,336 untuk program studi Teknik Informatika dan 0,446 untuk program studi Sistem Informasi. Berdasarkan analisis pemodelan topik, ditemukan bahwa topik-topik skripsi memiliki fokus yang jelas sesuai dengan visi dan misi masing-masing program studi.

Kata kunci: *Stemming; Skor koherensi; Judul skripsi; Visi misi; Pemodelan topik*

1. Pendahuluan

Skripsi merupakan hal paling penting bagi mahasiswa S1 yang wajib dilaksanakan agar dapat lulus dari institusi tempat kuliahnya. Skripsi menurut KBBI adalah karangan ilmiah yang wajib ditulis oleh mahasiswa sebagai bagian dari persyaratan akhir pendidikan akademisnya [1]. Topik skripsi pada perguruan tinggi secara umum harus sesuai dengan visi dan misi program studi agar memiliki keselarasan akademik, relevansi, dan kontribusi terhadap tujuan pendidikan institusi tersebut. Kegiatan penelitian di perguruan tinggi bertujuan untuk meningkatkan pengetahuan dan kemampuan perguruan tinggi sehingga bisa mencapai visi dan misi yang telah ditetapkan [2].

Perguruan tinggi XYZ yang menyelenggarakan pendidikan tinggi di bidang informatika dan komputer mempunyai 2 program studi, yaitu Teknik Informatika dan Sistem Informasi. Topik skripsi mahasiswa yang terdapat pada perguruan tinggi XYZ juga harus sesuai dengan visi dan misi program studi. Pada program studi Teknik Informatika memiliki visi penerapan sistem cerdas dan jaringan komputer nirkabel. Sedangkan pada program studi Sistem Informasi memiliki visi perancangan sistem informasi dan pemrograman berbasis manajemen bisnis. Namun, judul-judul skripsi yang telah ada, belum pernah dilakukan analisis kesesuaian topik pada judul-judul skripsi tersebut dengan visi masing-masing program studi. Sehingga belum dapat dipastikan apakah topik skripsi yang diambil oleh mahasiswa sudah sejalan dengan visi misi atau belum.

Oleh karena itu perlu dilakukan analisis topik berdasarkan judul skripsi agar dapat memastikan kesesuaian topik judul skripsi dengan visi program studi. Namun, dengan data yang

sangat banyak akan sulit untuk melakukan analisis dengan cepat dan tepat. Maka teknik yang dapat diterapkan untuk memproses analisis topik dengan cepat dan tepat adalah dengan teknik *Text Mining* yang disebut dengan pemodelan topik atau *Topic Modeling*. Penggunaan *Text Mining* dalam penelitian mendapatkan perhatian yang signifikan di berbagai aplikasi baik di dunia akademis maupun industri [3].

Pemodelan topik adalah pembentukan pola tertentu yang berasal dari sekumpulan teks [4]. Teknik pemodelan topik dapat secara efektif membagi sejumlah besar teks menjadi subruang topikal, yang membantu pemahaman dan analisis kumpulan teks yang sangat besar [5]. Salahsatu model yang diterapkan dalam pemodelan topik adalah *Latent Dirichlet Allocation* (LDA). LDA adalah model statistik generatif yang bertujuan untuk menentukan probabilitas berbagai topik yang mungkin ada dalam sebuah dokumen [6].

2. Tinjauan Pustaka

Beberapa penelitian terdahulu yang memiliki relevansi dengan penelitian ini seperti penelitian oleh [7] yang membahas pemodelan topik untuk mengetahui minat judul tugas akhir sebagai referensi mahasiswa dalam penentuan judul. Pemodelan topik dibuat menggunakan metode *Latent Dirichlet Allocation* (LDA) kemudian dibuat pengklasteran serta dilakukan evaluasi model menggunakan *hyperparameter tuning* dan *coherence score* sebagai alat ukur performa model. Setelah dilakukan *hyperparameter tuning* didapat *coherence score* sebesar 0,6177. Metode *hyperparameter tuning* berhubungan dengan cara mengambil sampel calon arsitektur model yang mungkin dari ruang nilai *hyperparameter* yang mungkin [8]. *Coherence score* atau Skor koherensi digunakan untuk menilai kualitas suatu topik dengan mengukur tingkat kesamaan semantik di antara kata-kata yang membentuk topik tersebut[9].

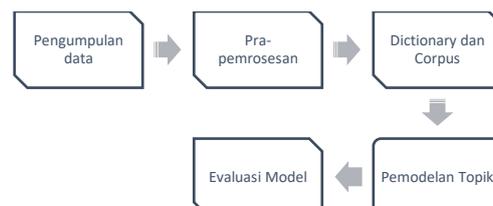
Penelitian yang dilakukan oleh [10] yang membahas Analisis Topik Penelitian Kesehatan di Indonesia Menggunakan Metode *Topic Modeling* LDA. Penelitian ini bertujuan untuk mengevaluasi kemampuan metode topic modeling LDA dalam menganalisis topik dari judul-judul penelitian di bidang kesehatan di Indonesia yang diambil dari Jurnal SINTA, serta untuk menilai skor koherensi setiap topik pada model yang dihasilkan. Hasil pengujian penelitian ini menggunakan kuisioner yang disampaikan responden peneliti, tenaga kesehatan, dan akademis rata-rata menjawab dengan pemodelan topik yang dihasilkan yaitu 94,1% mengatakan sangat baik dan 5,9 % baik.

Penelitian lain yaitu oleh [11] menganalisis tren topik penelitian tesis pada program studi magister ilmu komputer. Tujuannya adalah memetakan bidang-bidang penelitian yang telah dikerjakan serta mengidentifikasi bidang yang masih membutuhkan eksplorasi lebih lanjut. Metode yang digunakan adalah Latent Dirichlet Allocation (LDA) dengan data berupa 237 judul tesis, yang setelah diseleksi menghasilkan 211 judul. Tahapan preprocessing mencakup case folding, tokenizing, dan stopword removal tanpa penerapan stemming. Hasil penelitian menunjukkan nilai passes (jumlah iterasi) terbaik sebesar 28 dengan jumlah topik optimal sebanyak 5 topik. Dalam pemodelan topik, jumlah topik penting untuk menemukan kumpulan parameter yang optimal [12].

Penelitian [7] [10] [11] tidak menerapkan proses *Stemming* pada tahap *preprocessing* terhadap dataset judul penelitian. Berbeda dengan beberapa penelitian yang telah disebutkan di atas, penelitian ini membandingkan hasil pemodelan antara hasil dengan menerapkan *Stemming* dan tanpa *Stemming* pada tahap *preprocessing*.

3. Metodologi

Tahapan-tahapan yang dilakukan Pada penelitian ini, yaitu pengumpulan data, pra-pemrosesan (*casefolding, cleaning, stemming, stopword, tokenizing*), pembobotan dan pengujian algoritma.



Gambar 1. Tahapan penelitian

3.1. Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data judul skripsi 6 (enam) tahun terakhir pada program studi Teknik informatika dan Sistem Informasi, yaitu mulai tahun 2018 sampai dengan tahun 2024 sebanyak 1.086 data. Data sebanyak 759 dari program studi Teknik Informatika dan data dari program studi Sistem Informasi sebanyak 327 data.

3.2. Pra-pemrosesan Data

Pra-pemrosesan data merupakan serangkaian langkah yang dilakukan untuk mempersiapkan data sebelum dianalisis. Metode pra-pemrosesan memiliki pengaruh yang signifikan terhadap efektivitas klasifikasi teks pendek, di mana pendekatan sistematis yang mengintegrasikan berbagai teknik dapat meningkatkan kinerja sekaligus mempertahankan kelengkapan informasi [13].

Langkah-langkah tersebut yaitu:

1) Case Folding

Case folding merupakan langkah pra-pemrosesan data dengan mengubah teks dataset menjadi huruf kecil (*lower case*).

2) Cleaning Text

Cleaning text bertujuan untuk menghilangkan karakter dan kalimat yang tidak diperlukan, seperti tanda baca, simbol, URL, emotikon, dan elemen serupa lainnya. Proses ini dilakukan untuk memastikan teks menjadi lebih rapi dan terorganisir.

3) Tokenisasi

Tokenisasi merupakan langkah awal pemrosesan dengan melakukan pemecahan teks menjadi unit-unit dasar yang lebih kecil, yang disebut token. Tokenisasi yang umum digunakan yaitu memisahkan setiap kata dalam kalimat menggunakan tanda koma [14].

4) *Stopword Removal*

Stopword Removal merupakan langkah pra-pemrosesan data dengan menghapus kata sambung dan kata yang tidak efektif pada kalimat.

5) Lemmatization/*Stemming*

Fungsi lemmatization atau stemming adalah untuk mengubah kata-kata menjadi bentuk dasar atau akarnya (*stem*). *Stemming* dalam aplikasi NLP mengurangi volume teks, meningkatkan efisiensi dalam tugas-tugas seperti pengambilan informasi, klasifikasi teks, dan pengelompokan teks [15]. Teknik pengujian yang dilakukan pada penelitian ini, yaitu dengan menggunakan proses *stemming* dan tanpa menggunakan proses *stemming*.

6) *Vectorization*

Tujuan dari *vectorization* adalah membuat data mentah seperti teks menjadi format yang dapat diproses oleh model pembelajaran mesin. *Vectorization* text adalah konsep penting dalam tugas pemrosesan bahasa alami. Karena data teks tidak dapat langsung digunakan untuk pelatihan parameter model, maka perlu dilakukan *vectorization* data teks asli dan menjadikannya numerik, kemudian operasi ekstraksi fitur dapat dilakukan [16].

3.3. Pembuatan *Dictionary* dan *Corpus*

Langkah penting dalam pemrosesan teks yang mempersiapkan data untuk analisis oleh algoritma pembelajaran mesin adalah pembuatan *dictionary* dan *corpus*. *Dictionary* menyediakan struktur dan identifikasi kata-kata yang berbeda, dan *corpus* menyediakan representasi numerik dari teks yang dapat diolah oleh algoritma. Pada penelitian ini *dictionary* dan *corpus* dibuat menggunakan *library* gensim dari pemrograman python.

3.4. Pemodelan Topik Menggunakan LDA

Penerapan model LDA pada penelitian ini bertujuan untuk menganalisis judul-judul penelitian, dalam hal ini judul skripsi. *Latent Dirichlet Allocation* (LDA) adalah model generatif probabilistik yang digunakan untuk sekelompok data diskrit seperti kumpulan dokumen. LDA mengungkap topik laten (tersembunyi) dengan merepresentasikan setiap dokumen sebagai distribusi probabilitas dari beberapa topik, sementara setiap topik direpresentasikan sebagai distribusi probabilitas dari sejumlah kata [17].

Langkah selanjutnya adalah mengubah teks menjadi matriks fitur menggunakan Count Vectorizer, di mana setiap dokumen direpresentasikan sebagai baris dan setiap kata unik sebagai kolom. Kemudian, mendefinisikan model LDA dengan menentukan jumlah topik yang diinginkan

serta parameter lainnya seperti metode pembelajaran, jumlah iterasi maksimum, dan random state, lalu menerapkan model LDA pada matriks fitur tersebut.

Langkah berikutnya adalah tuning hyperparameter, yang dilakukan dengan teknik pencarian grid (*Grid Search*) untuk menemukan kombinasi hyperparameter terbaik. Hyperparameter yang diterapkan yaitu jumlah topik (5, 10, 15, 20), *learning_decay* (0.5, 0.7, 0.9), *max_iter* (10, 20, 30), *doc_topic_prior* (0.1, 0.5, 1.0) dan *topic_word_prior* (0.01, 0.1, 1.0). Kemudian cross-validation menggunakan KFold cross-validator untuk mengevaluasi performa model dengan berbagai kombinasi hyperparameter dan memilih matriks evaluasi yang sesuai, seperti *coherence score*, untuk mendapatkan set hyperparameter terbaik.

3.5. Evaluasi Model Menggunakan *Coherence Score*

Coherence score adalah metrik penting untuk mengevaluasi kualitas topik yang dihasilkan oleh model topik seperti LDA karena mengukur sejauh mana kata-kata dalam sebuah topik saling berkaitan, yang berkontribusi terhadap interpretabilitas dan relevansi topik tersebut. Skor ini membantu dalam memilih model terbaik melalui perbandingan berbagai konfigurasi parameter dan meningkatkan kualitas model dengan menghasilkan topik yang bermakna dan konsisten dengan evaluasi manual. Dengan menggunakan *coherence score*, kita dapat memastikan bahwa model topik tidak hanya memodelkan data secara statistik tetapi juga memberikan hasil yang bermanfaat dan mudah dipahami oleh pengguna.

Jenis *coherence score* yang digunakan pada penelitian ini adalah *C_v Coherence*. Metrik *C_v* mengevaluasi kualitas topik dengan menggabungkan konteks lokal kata-kata teratas, mengukur seberapa sering kata-kata tersebut muncul bersama (menggunakan NPMI), dan menghitung kesamaan mereka dalam ruang vektor (menggunakan cosine similarity) [18].

4. Hasil dan Pembahasan

4.1. Deskripsi Data

Data yang digunakan dalam penelitian ini adalah data judul skripsi 6 (enam) tahun terakhir pada program studi Teknik informatika dan Sistem Informasi, yaitu mulai tahun 2018 sampai dengan tahun 2024. Data judul skripsi Program Studi Teknik Informatika yang dikumpulkan sebagaimana pada tabel berikut:

Tabel 1. Dataset Judul Prodi TI

No.	Tahun	Jumlah
1	2018	133
2	2019	126
3	2020	135
4	2021	149
5	2022	79
6	2023	77
7	2024	60
Total		759

Kemudian data judul skripsi Program Studi Sistem Informasi yang dikumpulkan sebagaimana pada tabel berikut:

Tabel 2. Dataset Judul Prodi SI

No.	Tahun	Jumlah
1	2018	58
2	2019	63
3	2020	55
4	2021	62
5	2022	24
6	2023	41
7	2024	24
Total		327

Sebelum dilakukan proses pemodelan topik dengan LDA, data-data yang telah dikumpulkan akan dilakukan *preprocessing* terlebih dahulu.

4.2. Preprocessing (Pra-pemrosesan)

Untuk menjaga konsistensi dan kualitas data, berbagai tahap preprocessing dilakukan sebelum menerapkan pemodelan topik menggunakan LDA. Langkah pertama melibatkan normalisasi teks, di mana seluruh teks diubah menjadi huruf kecil untuk memastikan konsistensi. Kemudian, teks dibersihkan melalui penghapusan tanda baca, karakter khusus, angka, dan kata-kata yang tidak relevan seperti stopwords serta kata-kata dengan frekuensi kemunculan yang sangat rendah atau terlalu tinggi. Selanjutnya, proses lemmatisasi diterapkan untuk mengubah kata-kata ke dalam bentuk dasarnya. Tahapan ini dapat mencakup atau mengecualikan penggunaan algoritma stemming. Jika stemming digunakan, kata-kata akan diubah ke bentuk akarnya, misalnya "pembelajaran" menjadi "belajar." Sebaliknya, tanpa stemming, kata-kata dibiarkan dalam bentuk aslinya. Tahap terakhir adalah tokenisasi, yang memisahkan setiap kalimat dalam judul skripsi menjadi kata-kata individu. Sampel data yang telah melalui pembersihan teks dengan model dapat dilihat sebagai berikut:

1) Hasil pembersihan teks dengan *Stemming*

Tabel 3. Pembersihan teks Judul Teknik Informatika

No	Teks
1	sistem dukung putus pilih lokasi tps ahp fuzzy...
2	aplikasi layan administrasi bas web desa teluk...
3	aplikasi tata kelola uang desa studi desa simp...
4	sistem informasi dinas rumah mukim kota banjar...
5	sistem informasi administrasi monitoring sakit...
...	...
755	prototype pakan otomatis terna ikan program mi...
756	sistem presensi guru kenal wajah bas raspberry...
757	aplikasi web arsip bas full stack kantor desa ...
758	sistem nonitoring suhu airph air keruh air kol...
759	media ajar huruf hijaiyah tilawati jilid jilid...

Tabel 4. Pembersihan teks Judul Sistem Informasi

No	Teks
1	sistem informasi terima barang koleksi museum ...
2	sistem informasi kelola artefak balai arkeolog...
3	aplikasi chatting bas android ptpln persero wi...
4	aplikasi gaji excelso qmall banjarbaru
5	aplikasi izin bangun menara telekomunikasi din...
...	...
323	aplikasi kelola informasi kendara dinas operas...
324	sistem informasi layan posyandu desa stagen ba...
325	sistem informasi bayar spp bas web madrasah ts...
326	aplikasi catat uang usaha rahma store
327	aplikasi monitoring hasil produksi usaha kelap...

2) Hasil pembersihan teks tanpa *Stemming*

Tabel 5. Pembersihan teks Judul Teknik Informatika

No	Teks
1	sistem pendukung keputusan pemilihan lokasi tp...
2	aplikasi pelayanan administrasi berbasis web d...
3	aplikasi tata kelola keuangan desa studi desa ...
4	sistem informasi dinas perumahan pemukiman kot...
5	sistem informasi administrasi monitoring penya...
...	...
755	prototype pemberi pakan otomatis peternakan ik...
756	sistem presensi guru pengenalan wajah berbasis...
757	aplikasi web arsip berbasis full stack kantor ...
758	sistem nonitoring suhu airph air kekeruhan air...
759	media pembelajaran huruf hijaiyah tilawati jil...

Tabel 6. Pembersihan teks Judul Sistem Informasi

No	Teks
1	sistem informasi penerimaan barang koleksi mus...
2	sistem informasi pengelolaan artefak balai ark...

No	Teks
3	aplikasi chatting berbasis android ptpln perse...
4	aplikasi penggajian excelso qmall banjarbaru
5	aplikasi perizinan membangun menara telekomuni...
...	...
323	aplikasi pengelolaan informasi kendaraan dinas...
324	sistem informasi pelayanan posyandu desa stage...
325	sistem informasi pembayaran spp berbasis web m...
326	aplikasi pencatatan keuangan usaha rahma store
327	aplikasi monitoring hasil produksi usaha kelap...

Setelah proses pembersihan teks selesai, CountVectorizer digunakan untuk mengonversi teks menjadi matriks istilah-dokumen. Matriks ini merepresentasikan setiap dokumen (judul skripsi) sebagai baris dan setiap kata sebagai kolom, dengan nilai pada setiap sel menunjukkan frekuensi kemunculan kata tersebut dalam dokumen.

4.3. Pemodelan Topik

4.3.1. Parameter Model dengan Stemming

1) Program Studi Teknik Informatika

Tabel 7. Parameter TI Stemming

No.	Parameter	Nilai
1	doc_topic_prior	0.1
2	learning_decay	0.5
3	max_iter	30
4	n_components	5
5	topic_word_prior	1.0

Doc_topic_prior (0.1): Parameter ini menunjukkan bahwa distribusi topik dalam setiap dokumen cenderung lebih fokus atau terkonsentrasi, dengan lebih sedikit topik yang mungkin dikaitkan dengan setiap dokumen. **learning_decay (0.5):** Nilai 0.5 menunjukkan bahwa model akan menyesuaikan diri dengan perubahan data dengan moderat, menjaga keseimbangan antara adaptasi cepat dan stabilitas model. **max_iter (30):** Algoritma akan melakukan hingga 30 iterasi untuk mencapai konvergensi, memperbolehkan proses pembelajaran yang lebih dalam untuk meningkatkan akurasi model. **n_components (5):** Parameter ini menunjukkan bahwa model akan mengidentifikasi 5 topik utama dari dokumen yang dianalisis, mencerminkan kompleksitas atau variasi topik yang lebih rendah dalam data. **topic_word_prior (1.0):** Memberikan prioritas yang sama untuk semua kata dalam setiap topik, sehingga setiap kata memiliki kesempatan yang sama untuk muncul dalam topik yang ditemukan.

Tabel 8. Hasil pemodelan topik Parameter TI Stemming

No.	Topik
1	(0, '0.011*metode" + 0.007*saw" + 0.006*simple" + 0.006*weighting" + 0.006*additive" + 0.005*sistem" + 0.004*putus" + 0.004*tunjang" + 0.004*pilih" + 0.004*bola')
2	(1, '0.050*bas" + 0.028*sistem" + 0.018*informasi" + 0.017*aplikasi" + 0.016*web" + 0.015*arduino" + 0.015*prototype" + 0.012*banjarbaru" + 0.012*otomatis" + 0.009*android')
3	(2, '0.006*aplikasi" + 0.005*bas" + 0.005*game" + 0.004*edukasi" + 0.004*kenal" + 0.004*android" + 0.004*negeri" + 0.003*tk" + 0.002*guru" + 0.002*banjarbaru')
4	(3, '0.025*bas" + 0.025*metode" + 0.024*aplikasi" + 0.021*sistem" + 0.020*web" + 0.015*putus" + 0.013*pilih" + 0.010*banjarbaru" + 0.009*tunjang" + 0.007*additive')
5	(4, '0.029*metode" + 0.020*sistem" + 0.018*pakar" + 0.017*sakit" + 0.013*diagnosa" + 0.013*aplikasi" + 0.013*bas" + 0.010*prediksi" + 0.009*factor" + 0.008*certainty')

Coherence Score: 0.3273775141305622

2) Program Studi Sistem Informasi

Tabel 9. Parameter SI *Stemming*

No.	Parameter	Nilai
1	doc_topic_prior	1.0
2	learning_decay	0.5
3	max_iter	10
4	n_components	10
5	topic_word_prior	1.0

doc_topic_prior (1.0): Parameter ini menunjukkan bahwa setiap dokumen memiliki distribusi topik yang merata, artinya dokumen dapat terkait dengan banyak topik secara seimbang. **learning_decay (0.5):** Nilai 0.5 mengindikasikan bahwa model akan menyesuaikan diri dengan data baru pada tingkat yang moderat, menggabungkan adaptasi cepat dan stabilitas model. **max_iter (10):** Algoritma akan melakukan hingga 10 iterasi untuk menyempurnakan model sebelum berhenti, dengan maksimum upaya konvergensi yang ditetapkan. **n_components (10):** Menunjukkan bahwa model akan mengidentifikasi 10 topik utama dari dokumen yang dianalisis, sesuai dengan jumlah komponen yang ditentukan. **topic_word_prior (1.0):** Memberikan prioritas yang sama untuk semua kata dalam setiap topik, memastikan distribusi kata yang seragam dan tidak bias dalam pembentukan topik.

Tabel 10. Hasil pemodelan topik Parameter SI *Stemming*

No.	Topik
1	(0, '0.009*"aplikasi" + 0.009*"sistem" + 0.006*"informasi" + 0.006*"bas" + 0.004*"web" + 0.003*"banjarbaru" + 0.003*"kelola" + 0.003*"layan" + 0.002*"data" + 0.002*"selatan"')
2	(1, '0.009*"sistem" + 0.008*"aplikasi" + 0.006*"informasi" + 0.005*"bas" + 0.005*"web" + 0.003*"banjarbaru" + 0.003*"kelola" + 0.002*"layan" + 0.002*"dinas" + 0.002*"selatan"')
3	(2, '0.018*"aplikasi" + 0.011*"sistem" + 0.009*"bas" + 0.009*"web" + 0.008*"banjarbaru" + 0.008*"informasi" + 0.005*"data" + 0.005*"kelola" + 0.003*"kota" + 0.003*"selatan"')
4	(3, '0.013*"sistem" + 0.012*"metode" + 0.010*"aplikasi" + 0.007*"putus" + 0.006*"bas" + 0.005*"informasi" + 0.005*"pilih" + 0.005*"additive" + 0.005*"simple" + 0.005*"web"')
5	(4, '0.009*"sistem" + 0.009*"aplikasi" + 0.008*"informasi" + 0.008*"web" + 0.008*"bas" + 0.004*"kelola" + 0.003*"banjarbaru" + 0.003*"layan" + 0.003*"selatan" + 0.002*"data"')
6	(5, '0.044*"bas" + 0.035*"web" + 0.028*"aplikasi" + 0.027*"informasi" + 0.021*"sistem" + 0.009*"layan" + 0.007*"dinas" + 0.007*"banjarbaru" + 0.007*"kelola" + 0.006*"kabupaten"')
7	(6, '0.009*"sistem" + 0.007*"aplikasi" + 0.007*"bas" + 0.006*"web" + 0.005*"informasi" + 0.003*"banjarbaru" + 0.003*"kelola" + 0.002*"metode" + 0.002*"jual" + 0.002*"selatan"')
8	(7, '0.011*"aplikasi" + 0.008*"sistem" + 0.007*"bas" + 0.005*"web" + 0.005*"informasi" + 0.004*"kelola" + 0.004*"banjarbaru" + 0.003*"selatan" + 0.003*"kalimantan" + 0.002*"data"')
9	(8, '0.012*"sistem" + 0.010*"banjarbaru" + 0.010*"bas" + 0.009*"informasi" + 0.008*"aplikasi" + 0.007*"web" + 0.004*"kota" + 0.003*"layan" + 0.003*"data" + 0.002*"kelola"')
10	(9, '0.016*"aplikasi" + 0.008*"web" + 0.007*"bas" + 0.006*"sistem" + 0.006*"informasi" + 0.005*"banjarbaru" + 0.004*"kelola" + 0.003*"layan" + 0.003*"data" + 0.002*"jual"')

Coherence Score: 0.3287937509308532

4.3.2. Parameter Model tanpa *Stemming*

1) Program Studi Teknik Informatika

Tabel 11. Parameter TI Tanpa *Stemming*

No.	Parameter	Nilai
1	doc_topic_prior	0.1
2	learning_decay	0.5

3	max_iter	30
4	n_components	5
5	topic_word_prior	1.0

doc_topic_prior (0.1): Parameter ini menetapkan prior yang rendah untuk distribusi topik dalam setiap dokumen, yang berarti dokumen cenderung memiliki keterkaitan yang lebih fokus dengan beberapa topik tertentu. **learning_decay (0.5):** Dengan nilai 0.5, model akan menyesuaikan diri terhadap perubahan data dengan tingkat adaptasi yang moderat, mempertahankan stabilitas model sambil memungkinkan adaptasi terhadap data baru. **max_iter (30):** Ini adalah jumlah maksimum iterasi yang diizinkan untuk algoritma mencapai konvergensi, memungkinkan proses iterasi yang lebih panjang untuk meningkatkan kualitas model. **n_components (5):** Parameter ini menunjukkan bahwa model akan mencari 5 topik utama dalam data, mencerminkan kompleksitas atau variasi topik yang lebih rendah dalam analisis topik. **topic_word_prior (1.0):** Nilai ini menetapkan prior yang sama untuk semua kata dalam setiap topik, sehingga setiap kata memiliki kesempatan yang sama untuk muncul dalam setiap topik yang diidentifikasi oleh model.

Tabel 12. Hasil pemodelan topik Parameter TI Tanpa *Stemming*

No.	Topik
1	(0, '0.036*"metode" + 0.021*"sistem" + 0.019*"keputusan" + 0.017*"pemilihan" + 0.012*"penunjang" + 0.010*"additive" + 0.010*"simple" + 0.010*"saw" + 0.010*"weighting" + 0.007*"berbasis"')
2	(1, '0.046*"berbasis" + 0.023*"aplikasi" + 0.018*"sistem" + 0.014*"arduino" + 0.014*"web" + 0.014*"prototype" + 0.014*"banjarbaru" + 0.012*"android" + 0.011*"otomatis" + 0.009*"informasi"')
3	(2, '0.026*"berbasis" + 0.019*"aplikasi" + 0.017*"sistem" + 0.015*"informasi" + 0.014*"web" + 0.006*"android" + 0.006*"penjualan" + 0.005*"pemesanan" + 0.004*"kabupaten" + 0.004*"negeri"')
4	(3, '0.004*"metode" + 0.003*"prediksi" + 0.002*"rancang" + 0.002*"bangun" + 0.002*"game" + 0.002*"exponential" + 0.002*"implementasi" + 0.002*"mahasiswa" + 0.002*"robot" + 0.002*"rpg"')
5	(4, '0.024*"sistem" + 0.020*"metode" + 0.019*"pakar" + 0.018*"penyakit" + 0.015*"berbasis" + 0.013*"diagnosa" + 0.011*"web" + 0.010*"aplikasi" + 0.009*"factor" + 0.009*"certainty"')

Coherence Score: 0.3357855605057165

2) Program Studi Sistem Informasi

Tabel 13. Parameter SI Tanpa *Stemming*

No.	Parameter	Nilai
1	doc_topic_prior	1.0
2	learning_decay	0.5
3	max_iter	30
4	n_components	10
5	topic_word_prior	1.0

doc_topic_prior (1.0): Parameter ini menetapkan prior yang tinggi untuk distribusi topik dalam setiap dokumen. Dengan nilai 1.0, setiap dokumen diharapkan memiliki distribusi topik yang merata atau lebih tersebar, memungkinkan setiap dokumen terkait dengan banyak topik. **learning_decay (0.5):** Nilai 0.5 mengontrol laju pembelajaran model. Ini menyeimbangkan antara adaptasi cepat terhadap perubahan data baru dan stabilitas model. Model akan menyesuaikan diri dengan data baru pada tingkat yang moderat. **max_iter (30):** Parameter ini menentukan jumlah maksimum iterasi yang diizinkan sebelum algoritma berhenti. Dengan nilai 30, algoritma akan melakukan hingga 30 iterasi untuk mencoba mencapai konvergensi atau kondisi di mana model tidak berubah lagi secara signifikan. **n_components (10):** Ini menunjukkan jumlah topik yang akan dicari dalam data. Dengan nilai 10, model akan mengidentifikasi dan memodelkan 10 topik utama dari koleksi dokumen yang diberikan. **topic_word_prior (1.0):** Parameter ini menetapkan prior yang sama untuk semua kata dalam setiap topik. Nilai 1.0 menunjukkan bahwa setiap kata memiliki kesempatan yang sama untuk muncul dalam setiap topik yang diidentifikasi oleh model.

Tabel 14. Hasil pemodelan topik Parameter SI Tanpa *Stemming*

No.	Topik
1	(0, '0.002*"aplikasi" + 0.001*"banjarmasin" + 0.001*"pt" + 0.001*"penjualan" + 0.001*"sistem" + 0.001*"manajemen" + 0.001*"puskesmas" + 0.001*"android" + 0.001*"toko" + 0.001*"ikan"')
2	(1, '0.003*"aplikasi" + 0.001*"pt" + 0.001*"banjarmasin" + 0.001*"penjualan" + 0.001*"sistem" + 0.001*"manajemen" + 0.001*"puskesmas" + 0.001*"android" + 0.001*"toko" + 0.001*"berbasis"')
3	(2, '0.003*"aplikasi" + 0.001*"penjualan" + 0.001*"banjarmasin" + 0.001*"pt" + 0.001*"sistem" + 0.001*"manajemen" + 0.001*"puskesmas" + 0.001*"toko" + 0.001*"android" + 0.001*"ikan"')
4	(3, '0.002*"aplikasi" + 0.001*"penjualan" + 0.001*"sistem" + 0.001*"banjarmasin" + 0.001*"pt" + 0.001*"manajemen" + 0.001*"puskesmas" + 0.001*"android" + 0.001*"toko" + 0.001*"berbasis"')
5	(4, '0.051*"berbasis" + 0.050*"aplikasi" + 0.043*"sistem" + 0.043*"web" + 0.039*"informasi" + 0.020*"banjarbaru" + 0.013*"pengelolaan" + 0.009*"dinas" + 0.008*"selatan" + 0.008*"pelayanan"')
6	(5, '0.020*"metode" + 0.010*"keputusan" + 0.010*"sistem" + 0.009*"pemilihan" + 0.007*"additive" + 0.006*"simple" + 0.006*"saw" + 0.006*"weighting" + 0.005*"penunjang" + 0.004*"pendukung"')
7	(6, '0.003*"aplikasi" + 0.001*"banjarmasin" + 0.001*"pt" + 0.001*"penjualan" + 0.001*"sistem" + 0.001*"manajemen" + 0.001*"puskesmas" + 0.001*"android" + 0.001*"toko" + 0.001*"berbasis"')
8	(7, '0.003*"aplikasi" + 0.001*"pt" + 0.001*"banjarmasin" + 0.001*"penjualan" + 0.001*"sistem" + 0.001*"manajemen" + 0.001*"puskesmas" + 0.001*"android" + 0.001*"toko" + 0.001*"ikan"')
9	(8, '0.003*"aplikasi" + 0.002*"penjualan" + 0.001*"sistem" + 0.001*"banjarmasin" + 0.001*"pt" + 0.001*"manajemen" + 0.001*"toko" + 0.001*"puskesmas" + 0.001*"berbasis" + 0.001*"android"')
10	(9, '0.003*"aplikasi" + 0.002*"penjualan" + 0.002*"sistem" + 0.001*"motor" + 0.001*"berbasis" + 0.001*"banjarmasin" + 0.001*"pt" + 0.001*"android" + 0.001*"toko" + 0.001*"manajemen"')

Coherence Score: 0.445669348220113

4.4. Hasil Analisis

4.4.1. Topik Skripsi Teknik Informatika

1) Topik Menerapkan *Stemming*

Topik 1: Metode Pendukung Keputusan

Topik: (0, '0.011*"metode" + 0.007*"saw" + 0.006*"simple" + 0.006*"weighting" + 0.006*"additive" + 0.005*"sistem" + 0.004*"putus" + 0.004*"tunjang" + 0.004*"pilih" + 0.004*"bola"')

Topik 1 menekankan pada berbagai metode untuk mendukung keputusan, seperti Simple Additive Weighting (SAW) dan teknik weighting. Ini sangat relevan dengan visi program studi yang menekankan penerapan sistem cerdas, karena penggunaan metode-metode ini adalah bagian dari pengembangan sistem yang pintar. Selain itu, ini mendukung misi program studi untuk menghasilkan sarjana yang profesional dalam penerapan sistem cerdas dan jaringan komputer nirkabel.

Topik 2: Aplikasi dan Sistem Berbasis Web dan IoT

Topik: (1, '0.050*"bas" + 0.028*"sistem" + 0.018*"informasi" + 0.017*"aplikasi" + 0.016*"web" + 0.015*"arduino" + 0.015*"prototype" + 0.012*"banjarbaru" + 0.012*"otomatis" + 0.009*"android"')

Topik 2 fokus pada pengembangan aplikasi dan sistem berbasis web, termasuk teknologi Internet of Things (IoT) seperti Arduino dan otomatisasi. Ini sangat sejalan dengan visi program studi yang mengutamakan penerapan jaringan komputer nirkabel dan sistem cerdas. Ini juga mendukung misi untuk menghasilkan penelitian dan publikasi di bidang teknik informatika, serta meningkatkan kapasitas dan reputasi program studi.

Topik 3: Aplikasi Game dan Edukasi

Topik: (2, '0.006*"aplikasi" + 0.005*"bas" + 0.005*"game" + 0.004*"edukasi" + 0.004*"kenal" + 0.004*"android" + 0.004*"negeri" + 0.003*"tk" + 0.002*"guru" + 0.002*"banjarbaru"')

Topik 3 menekankan pada pengembangan aplikasi game dan edukasi, yang mencakup platform Android. Fokus pada pengembangan game dan aplikasi edukasi menunjukkan kemampuan mahasiswa untuk menciptakan solusi kreatif yang relevan dengan kebutuhan

masyarakat, mendukung misi program studi dalam pengabdian masyarakat dan tanggung jawab sosial.

Topik 4: Metode dan Sistem Pendukung Keputusan

Topik: (3, '0.025**"bas" + 0.025**"metode" + 0.024**"aplikasi" + 0.021**"sistem" + 0.020**"web" + 0.015**"putus" + 0.013**"pilih" + 0.010**"banjarbaru" + 0.009**"tunjang" + 0.007**"additive"')

Topik 4 mencakup berbagai metode dan aplikasi sistem pendukung keputusan berbasis web. Ini sesuai dengan visi untuk menjadi unggulan dalam penerapan sistem cerdas dan jaringan komputer nirkabel. Pengembangan sistem pendukung keputusan mencerminkan misi untuk mengembangkan pendidikan dan kurikulum yang bermutu, serta menghasilkan karya ilmiah bermutu di bidang teknik informatika.

Topik 5: Sistem Pakar dan Diagnosa Penyakit

Topik: (4, '0.029**"metode" + 0.020**"sistem" + 0.018**"pakar" + 0.017**"sakit" + 0.013**"diagnosa" + 0.013**"aplikasi" + 0.013**"bas" + 0.010**"prediksi" + 0.009**"factor" + 0.008**"certainty"')

Topik 5 berfokus pada sistem pakar dan diagnosa penyakit, menggunakan metode tertentu untuk diagnosa. Ini sangat relevan dengan visi untuk mengembangkan sistem cerdas, serta misi untuk melaksanakan dan mempublikasikan penelitian di bidang teknik informatika. Sistem pakar untuk diagnosa penyakit menunjukkan kontribusi langsung terhadap peningkatan kesejahteraan masyarakat, sejalan dengan tujuan program studi.

Pemodelan topik dengan *stemming* ini menunjukkan bahwa judul-judul skripsi dari Program Studi Teknik Informatika mencakup berbagai aspek yang mendukung visi dan misi program studi. Topik-topik yang dihasilkan mencerminkan fokus pada penerapan sistem cerdas, jaringan komputer nirkabel, dan pengembangan aplikasi berbasis teknologi terbaru. Hal ini menunjukkan bahwa penelitian yang dilakukan oleh mahasiswa sudah sejalan dengan tujuan untuk menghasilkan lulusan berkualitas dan karya ilmiah yang bermutu tinggi, serta mengembangkan kerjasama dengan berbagai institusi untuk meningkatkan reputasi program studi.

2) *Topik Tanpa Stemming*

Topik 1: Metode dan Sistem Keputusan

Topik: (0, '0.036**"metode" + 0.021**"sistem" + 0.019**"keputusan" + 0.017**"pemilihan" + 0.012**"penunjang" + 0.010**"additive" + 0.010**"simple" + 0.010**"saw" + 0.010**"weighting" + 0.007**"berbasis"')

Topik 1 menggarisbawahi penggunaan berbagai metode dalam sistem pendukung keputusan, seperti SAW (Simple Additive Weighting) dan teknik-teknik berbasis web. Fokus pada metode, pemilihan, dan penunjang keputusan ini relevan dengan visi untuk menjadi program studi unggulan dalam penerapan sistem cerdas, serta misi untuk mengembangkan pendidikan yang bermutu tinggi di bidang teknik informatika.

Topik 2: Aplikasi Berbasis Web dan IoT

Topik: (1, '0.046**"berbasis" + 0.023**"aplikasi" + 0.018**"sistem" + 0.014**"arduino" + 0.014**"web" + 0.014**"prototype" + 0.014**"banjarbaru" + 0.012**"android" + 0.011**"otomatis" + 0.009**"informasi"')

Topik 2 berfokus pada pengembangan aplikasi berbasis web dan teknologi Internet of Things (IoT), termasuk penggunaan Arduino dan otomatisasi. Ini sangat sesuai dengan visi program studi yang menekankan pada penerapan jaringan komputer nirkabel dan sistem cerdas, serta misi untuk menghasilkan sarjana yang profesional dalam teknologi ini.

Topik 3: Aplikasi dan Sistem Informasi

Topik: (2, '0.026**"berbasis" + 0.019**"aplikasi" + 0.017**"sistem" + 0.015**"informasi" + 0.014**"web" + 0.006**"android" + 0.006**"penjualan" + 0.005**"pemesanan" + 0.004**"kabupaten" + 0.004**"negeri"')

Topik 3 mencakup pengembangan aplikasi dan sistem informasi yang berbasis web dan android untuk berbagai konteks, seperti penjualan dan pemesanan. Ini mendukung visi dan misi program studi dengan menunjukkan aplikasi praktis dan solusi inovatif dalam sistem informasi, yang merupakan bagian penting dari teknik informatika.

Topik 4: Metode Prediksi dan Rancang Bangun

Topik: (3, '0.004**"metode" + 0.003**"prediksi" + 0.002**"rancang" + 0.002**"bangun" + 0.002**"game" + 0.002**"exponential" + 0.002**"implementasi" + 0.002**"mahasiswa" + 0.002**"robot" + 0.002**"rpg"')

Topik 4 menekankan pada metode prediksi, perancangan, dan pembangunan sistem, termasuk pengembangan game dan robot. Ini relevan dengan misi program studi untuk mengembangkan kurikulum yang bermutu tinggi dan menghasilkan karya ilmiah yang bermutu di

bidang teknik informatika. Fokus pada implementasi dan rancang bangun juga mencerminkan tujuan menghasilkan lulusan yang kompeten dalam menghadapi tantangan global.

Topik 5: Sistem Pakar dan Diagnosa Penyakit

Topik: (4, '0.024*"sistem" + 0.020*"metode" + 0.019*"pakar" + 0.018*"penyakit" + 0.015*"berbasis" + 0.013*"diagnosa" + 0.011*"web" + 0.010*"aplikasi" + 0.009*"factor" + 0.009*"certainty"')

Topik 5 berfokus pada sistem pakar dan diagnosa penyakit, yang merupakan aplikasi penting dari sistem cerdas. Ini mendukung visi program studi untuk menjadi unggulan dalam penerapan sistem cerdas dan menunjukkan misi untuk melaksanakan penelitian yang dapat dipublikasikan. Penggunaan sistem pakar dalam diagnosa penyakit juga mencerminkan tanggung jawab sosial untuk mengembangkan aplikasi yang relevan bagi kebutuhan masyarakat.

Pemodelan topik tanpa *stemming* ini menunjukkan bahwa judul-judul skripsi dari Program Studi Teknik Informatika mencakup berbagai aspek yang mendukung visi dan misi program studi. Topik-topik yang dihasilkan mencerminkan fokus pada penerapan sistem cerdas, jaringan komputer nirkabel, dan pengembangan aplikasi berbasis teknologi terbaru. Hal ini menunjukkan bahwa penelitian yang dilakukan oleh mahasiswa sudah sejalan dengan tujuan untuk menghasilkan lulusan berkualitas dan karya ilmiah yang bermutu tinggi, serta mengembangkan kerjasama dengan berbagai institusi untuk meningkatkan reputasi program studi.

3) *Visualisasi Topik Skripsi Teknik Informatika*

Dalam bentuk wordcloud Topik skripsi Teknik Informatika divisualisasikan sebagai berikut:



Gambar 2. Visualisasi Topik TI

4.4.2. Topik Skripsi Sistem Informasi

1) *Topik Menerapkan Stemming*

Topik 1: Pengembangan Aplikasi dan Sistem Informasi

Topik: (0, '0.009*"aplikasi" + 0.009*"sistem" + 0.006*"informasi" + 0.006*"bas" + 0.004*"web" + 0.003*"banjarbaru" + 0.003*"kelola" + 0.003*"layan" + 0.002*"data" + 0.002*"selatan"')

Topik 1 mencakup pengembangan aplikasi dan sistem informasi yang berbasis web dengan fokus pada pengelolaan layanan dan data, terutama di wilayah Banjarbaru dan Kalimantan Selatan. Ini sangat relevan dengan visi untuk menjadi unggulan dalam perancangan sistem informasi dan pemrograman berbasis manajemen bisnis. Pengelolaan layanan dan data sesuai dengan misi pengembangan teknologi informasi yang berbasis manajemen bisnis.

Topik 2: Sistem Informasi untuk Dinas

Topik: (1, '0.009*"sistem" + 0.008*"aplikasi" + 0.006*"informasi" + 0.005*"bas" + 0.005*"web" + 0.003*"banjarbaru" + 0.003*"kelola" + 0.002*"layan" + 0.002*"dinas" + 0.002*"selatan"')

Topik 2 berfokus pada sistem informasi yang digunakan dalam aplikasi berbasis web dan layanan untuk dinas di wilayah Banjarbaru dan Kalimantan Selatan. Fokus pada aplikasi dan sistem informasi untuk dinas sesuai dengan misi pengembangan sistem informasi yang bermanfaat bagi manajemen bisnis dan kebutuhan masyarakat.

Topik 3: Aplikasi Web dan Pengelolaan Data

Topik: (2, '0.018*"aplikasi" + 0.011*"sistem" + 0.009*"bas" + 0.009*"web" + 0.008*"banjarbaru" + 0.008*"informasi" + 0.005*"data" + 0.005*"kelola" + 0.003*"kota" + 0.003*"selatan"')

Topik 3 menekankan pada pengembangan aplikasi berbasis web dan sistem informasi untuk pengelolaan data dan layanan di kota Banjarbaru dan Kalimantan Selatan. Ini sesuai dengan visi untuk menghasilkan sarjana yang profesional dalam perancangan sistem informasi

dan pemrograman berbasis manajemen bisnis, serta pengembangan teknologi informasi untuk masyarakat.

Topik 4: Sistem Pendukung Keputusan

Topik: (3, '0.013**sistem" + 0.012**metode" + 0.010**aplikasi" + 0.007**putus" + 0.006**bas" + 0.005**informasi" + 0.005**pilih" + 0.005**additive" + 0.005**simple" + 0.005**web")

Topik 4 mencakup metode dan aplikasi untuk sistem pendukung keputusan, termasuk metode additive dan simple. Penggunaan metode dalam sistem pendukung keputusan sesuai dengan misi untuk mengembangkan penelitian di bidang teknologi sistem informasi.

Topik 5: Sistem Informasi dan Pengelolaan Layanan

Topik: (4, '0.009**sistem" + 0.009**aplikasi" + 0.008**informasi" + 0.008**web" + 0.008**bas" + 0.004**kelola" + 0.003**banjarbaru" + 0.003**layan" + 0.003**selatan" + 0.002**data")

Topik 5 berfokus pada pengembangan sistem informasi dan aplikasi web yang mengelola data dan layanan di wilayah Banjarbaru dan Kalimantan Selatan. Fokus pada pengelolaan data dan layanan sesuai dengan misi untuk mempublikasikan hasil penelitian dan mengembangkan teknologi informasi yang relevan bagi kebutuhan masyarakat.

Topik 6: Aplikasi Web dan Layanan untuk Dinas

Topik: (5, '0.044**bas" + 0.035**web" + 0.028**aplikasi" + 0.027**informasi" + 0.021**sistem" + 0.009**layan" + 0.007**dinas" + 0.007**banjarbaru" + 0.007**kelola" + 0.006**kabupaten")

Topik 6 berfokus pada aplikasi web dan sistem informasi untuk layanan dan pengelolaan di dinas dan kabupaten Banjarbaru. Topik ini mendukung misi untuk mengembangkan penelitian dan aplikasi teknologi informasi yang bermanfaat bagi institusi regional dan nasional.

Topik 7: Pengelolaan Data Penjualan

Topik: (6, '0.009**sistem" + 0.007**aplikasi" + 0.007**bas" + 0.006**web" + 0.005**informasi" + 0.003**banjarbaru" + 0.003**kelola" + 0.002**metode" + 0.002**jual" + 0.002**selatan")

Topik 7 mencakup pengembangan sistem informasi dan aplikasi web, termasuk metode untuk pengelolaan data penjualan di wilayah Banjarbaru dan Kalimantan Selatan. Topik ini relevan dengan visi dan misi untuk menghasilkan lulusan yang profesional dalam manajemen bisnis berbasis teknologi informasi.

Topik 8: Aplikasi Web untuk Pengelolaan Layanan

Topik: (7, '0.011**aplikasi" + 0.008**sistem" + 0.007**bas" + 0.005**web" + 0.005**informasi" + 0.004**kelola" + 0.004**banjarbaru" + 0.003**selatan" + 0.003**kalimantan" + 0.002**data")

Topik 8 berfokus pada aplikasi dan sistem informasi berbasis web untuk pengelolaan data dan layanan di Kalimantan Selatan. Fokus pada pengelolaan informasi dan layanan mendukung misi untuk mempublikasikan hasil penelitian yang bermanfaat bagi kebutuhan masyarakat.

Topik 9: Sistem Informasi untuk Kota Banjarbaru

Topik: (8, '0.012**sistem" + 0.010**banjarbaru" + 0.010**bas" + 0.009**informasi" + 0.008**aplikasi" + 0.007**web" + 0.004**kota" + 0.003**layan" + 0.003**data" + 0.002**kelola")

Topik 9 mencakup pengembangan sistem informasi untuk layanan dan pengelolaan data di kota Banjarbaru. Sesuai dengan misi untuk mengembangkan kegiatan pengabdian kepada masyarakat melalui teknologi informasi yang relevan.

Topik 10: Aplikasi Web untuk Layanan Penjualan

Topik: (9, '0.016**aplikasi" + 0.008**web" + 0.007**bas" + 0.006**sistem" + 0.006**informasi" + 0.005**banjarbaru" + 0.004**kelola" + 0.003**layan" + 0.003**data" + 0.002**jual")

Topik 10 berfokus pada aplikasi web dan sistem informasi untuk layanan dan pengelolaan data penjualan di Banjarbaru. Mendukung visi untuk menghasilkan sarjana yang profesional dalam perancangan sistem informasi dan manajemen bisnis.

Topik-topik yang dihasilkan sangat relevan dengan visi dan misi Program Studi Sistem Informasi. Fokus pada aplikasi web, sistem informasi, dan layanan pengelolaan sesuai dengan tujuan untuk menghasilkan lulusan yang mampu bersaing di bidang perancangan sistem informasi dan pemrograman berbasis manajemen bisnis.

2) *Topik Tanpa Stemming***Topik 1: Aplikasi dan Sistem Manajemen**

Topik: (0, '0.002**"aplikasi" + 0.001**"banjarmasin" + 0.001**"pt" + 0.001**"penjualan" + 0.001**"sistem" + 0.001**"manajemen" + 0.001**"puskesmas" + 0.001**"android" + 0.001**"toko" + 0.001**"ikan"')

Topik 1 menyoroti pengembangan aplikasi untuk mendukung berbagai keperluan manajemen, seperti manajemen penjualan dan manajemen di puskesmas. Hal ini sejalan dengan visi program studi yang berkomitmen pada keunggulan dalam perancangan sistem informasi dan pengembangan perangkat lunak berbasis manajemen bisnis. Selain itu, pengembangan aplikasi berbasis Android mencerminkan misi program studi dalam mempublikasikan penelitian di bidang teknologi informasi yang relevan dengan kebutuhan masyarakat.

Topik 2: Aplikasi Penjualan dan Sistem Manajemen

Topik: (1, '0.003**"aplikasi" + 0.001**"pt" + 0.001**"banjarmasin" + 0.001**"penjualan" + 0.001**"sistem" + 0.001**"manajemen" + 0.001**"puskesmas" + 0.001**"android" + 0.001**"toko" + 0.001**"berbasis"')

Topik 2 juga menekankan pada pengembangan aplikasi penjualan dan sistem manajemen untuk perusahaan di Banjarmasin. Fokus pada aplikasi berbasis manajemen bisnis mendukung visi untuk menghasilkan sarjana yang profesional dalam perancangan sistem informasi, serta misi untuk mengembangkan teknologi informasi yang berbasis manajemen bisnis.

Topik 3: Aplikasi dan Sistem Penjualan

Topik: (2, '0.003**"aplikasi" + 0.001**"penjualan" + 0.001**"banjarmasin" + 0.001**"pt" + 0.001**"sistem" + 0.001**"manajemen" + 0.001**"puskesmas" + 0.001**"toko" + 0.001**"android" + 0.001**"ikan"')

Topik 3 menyoroti pengembangan aplikasi untuk penjualan dan sistem manajemen, khususnya di sektor puskesmas dan toko ikan di Banjarmasin. Hal ini mendukung misi untuk melaksanakan penelitian di bidang teknologi sistem informasi yang berbasis teknologi informasi dan manajemen bisnis.

Topik 4: Aplikasi Penjualan dan Manajemen

Topik: (3, '0.002**"aplikasi" + 0.001**"penjualan" + 0.001**"sistem" + 0.001**"banjarmasin" + 0.001**"pt" + 0.001**"manajemen" + 0.001**"puskesmas" + 0.001**"android" + 0.001**"toko" + 0.001**"berbasis"')

Topik 4 mencakup pengembangan aplikasi penjualan dan sistem manajemen berbasis Android untuk berbagai sektor termasuk puskesmas dan toko. Fokus ini mendukung visi untuk menghasilkan lulusan yang kompeten dalam perancangan sistem informasi.

Topik 5: Aplikasi Berbasis Web untuk Pengelolaan Informasi

Topik: (4, '0.051**"berbasis" + 0.050**"aplikasi" + 0.043**"sistem" + 0.043**"web" + 0.039**"informasi" + 0.020**"banjarbaru" + 0.013**"pengelolaan" + 0.009**"dinas" + 0.008**"selatan" + 0.008**"pelayanan"')

Topik 5 berfokus pada pengembangan aplikasi berbasis web untuk pengelolaan informasi di berbagai dinas di Banjarbaru dan Kalimantan Selatan. Fokus pada aplikasi web mendukung visi untuk menjadi program studi unggulan dalam perancangan sistem informasi dan pemrograman berbasis manajemen bisnis. Topik ini juga sejalan dengan misi untuk mengembangkan penelitian dan teknologi informasi yang relevan bagi masyarakat.

Topik 6: Sistem Pendukung Keputusan

Topik: (5, '0.020**"metode" + 0.010**"keputusan" + 0.010**"sistem" + 0.009**"pemilihan" + 0.007**"additive" + 0.006**"simple" + 0.006**"saw" + 0.006**"weighting" + 0.005**"penunjang" + 0.004**"pendukung"')

Topik 6 menyoroti metode dan sistem pendukung keputusan, termasuk metode SAW (Simple Additive Weighting). Fokus ini mendukung misi untuk melaksanakan penelitian di bidang teknologi sistem informasi, serta mempublikasikan hasil penelitian tersebut.

Topik 7: Aplikasi Penjualan dan Sistem Manajemen

Topik: (6, '0.003**"aplikasi" + 0.001**"banjarmasin" + 0.001**"pt" + 0.001**"penjualan" + 0.001**"sistem" + 0.001**"manajemen" + 0.001**"puskesmas" + 0.001**"android" + 0.001**"toko" + 0.001**"berbasis"')

Topik 7 menekankan pada aplikasi penjualan dan sistem manajemen di berbagai sektor, termasuk puskesmas dan toko di Banjarmasin. Hal ini sesuai dengan misi untuk mengembangkan kurikulum bermutu tinggi dalam bidang sistem informasi yang berbasis manajemen bisnis.

Hasil analisis dalam penelitian ini menunjukkan bahwa penerapan model *Latent Dirichlet Allocation* (LDA) dalam pemodelan topik memberikan wawasan mendalam tentang tren dan pola dalam pemilihan judul penelitian, serta meningkatkan pemahaman mengenai relevansi penelitian di berbagai bidang. Penelitian ini sejalan dengan penelitian oleh [7], yang menunjukkan efektivitas LDA dalam menganalisis minat mahasiswa terhadap judul tugas akhir, di mana pengklasteran yang dilakukan menghasilkan nilai *coherence score* yang signifikan. Selain itu, temuan penelitian ini memperkuat penelitian oleh [10], yang mengidentifikasi bahwa 94,1% responden menilai hasil pemodelan topik dalam analisis kesehatan sebagai sangat baik, mencerminkan akurasi dan kehandalan metode LDA. Penelitian oleh [11] juga menyoroti pentingnya pemodelan topik dalam memetakan tren penelitian di program studi magister ilmu komputer, di mana LDA berhasil menghasilkan parameter optimal.

Pendekatan yang lebih terfokus dalam analisis judul tugas akhir mahasiswa, yang memungkinkan pengidentifikasian pola pemilihan judul yang lebih spesifik, serta penerapan evaluasi model yang lebih komprehensif dengan membandingkan berbagai metrik kinerja menjadi *Novelty* pada penelitian ini. Dengan demikian, penelitian ini tidak hanya menegaskan kembali efektivitas model LDA sebagai alat yang relevan dan bermanfaat dalam pemodelan topik di berbagai bidang, tetapi juga memberikan kontribusi baru dalam memahami dinamika akademis yang mendasari pemilihan judul penelitian.

5. Simpulan

Berdasarkan analisis pemodelan topik untuk program studi Teknik Informatika dan Sistem Informasi, ditemukan bahwa kedua program studi memiliki fokus yang jelas sesuai dengan visi dan misi mereka masing-masing. Meskipun terdapat 1 topik dari 2 program studi tersebut yang masih belum sesuai. Pemodelan topik dengan menggunakan teknik *stemming* dan tanpa *stemming* pada dasarnya menghasilkan topik-topik yang serupa, menunjukkan konsistensi dalam temuan terhadap area keahlian yang diutamakan oleh program studi. *Stemming* cenderung menyederhanakan kata kunci yang bisa mengurangi kejelasan konteks topik, namun dapat memperluas cakupan aplikasi dalam penelitian. Tanpa *stemming* umumnya menghasilkan *coherence score* yang lebih tinggi dibandingkan dengan *stemming*, menunjukkan bahwa informasi yang lebih lengkap dalam kata kunci membantu dalam membentuk topik yang lebih kohesif. Program studi Teknik Informatika terutama menyoroti bidang sistem pakar, keamanan jaringan, pengolahan citra, dan aplikasi web, sementara Sistem Informasi menekankan pada aplikasi manajemen, penjualan, dan pengambilan keputusan berbasis sistem informasi. Relevansi topik-topik ini dengan visi untuk unggul dalam teknologi informatika dan sistem informasi terbukti kuat, dengan potensi kontribusi yang signifikan terhadap pengembangan teknologi dan penerapan praktis di masyarakat.

Rekomendasi untuk penelitian di masa mendatang yaitu seperti memperhatikan teknik pra-proses yang lebih canggih selain *stemming*, seperti *lemmatization* dan penghapusan *stopwords* yang lebih efektif, untuk meningkatkan *coherence score* dan relevansi topik.

Daftar Referensi

- [1] KEMDIKBUDRISTEK RI, "Pencarian: KBBI Daring," KBBI Daring, 2016. [Online]. Available: <https://kbbi.kemdikbud.go.id>. [Diakses 22 Mei 2024].
- [2] J. Kim, "An Exploratory Study on the Meaning of Institutional Research (IR)," *The Korean Society for the Study of Local Education Management*, vol. 26, no. 3, pp. 55-81, 2023.
- [3] T. Muhayat, A. Fauzi dan J. Indra, "Analisis Sentimen Terhadap Komentar Video Youtube Menggunakan Support Vector Machines," *Progresif: Jurnal Ilmiah Komputer*, vol. 19, no. 1, pp. 231-240, 2023.
- [4] A. R. D. Astuti dan N. Cahyono, "Analisis Topic Modelling Persepsi Pengguna Internet Menggunakan Metode Latent Dirichlet Allocation," *Indonesian Journal of Computer Science*, vol. 12, no. 1, pp. 326-334, 2023.
- [5] U. Chauhan dan A. Shah, "Topic Modeling Using Latent Dirichlet allocation: A Survey," *ACM Computing Surveys (CSUR)*, vol. 54, no. 7, pp. 1-35, 2021.
- [6] S. Bhutada, *Topic Modeling Using Variations On Latent Dirichlet Allocation*, Raleigh: Lulu Publication, 2020.

- [7] T. Titiana dan D. H. Bangkalang, "Analisis Dan Penerapan Topic Modeling Pada Judul Tugas Akhir Mahasiswa Menggunakan Metode Latent Dirichlet Allocation (LDA)," *Jurnal Ilmiah Penelitian dan Pembelajaran Informatika (JIPI)*, vol. 8, no. 4, pp. 1275-1287, 2023.
- [8] T. A. E. Putri, T. Widiharih dan R. Santoso, "Penerapan Tuning Hyperparameter Randomsearchcv Pada Adaptive Boosting Untuk Prediksi Kelangsungan Hidup Pasien Gagal Jantung," *Jurnal Gaussian*, vol. 11, no. 3, pp. 397-406, 2022.
- [9] H. Koh dan M. Fienup, "Topic Modeling as a Tool for Analyzing Library Chat Transcripts," *Information Technology and Libraries*, vol. 40, no. 3, pp. 1-24, 2021.
- [10] Y. Sahria dan D. H. Fudholi, "Analisis Topik Penelitian Kesehatan di Indonesia Menggunakan Metode Topic Modeling LDA (Latent Dirichlet Allocation)," *Rekayasa Sistem dan Teknologi Informasi (RESTI)*, vol. 4, no. 2, pp. 336-344, 2020.
- [11] L. Bayuaji dan A. Wahyudi, "Analisis Trend Topik Penelitian Tesis Pada Program Studi Magister Ilmu Komputer Universitas Budi Luhur Menggunakan Metode Latent Dirichlet Allocation (LDA)," *Faktor Exacta*, vol. 17, no. 1, pp. 77-84, 2024.
- [12] S. Mifrah dan E. H. Benlahmar, "Topic Modeling *Coherence*: A Comparative Study between LDA and NMF Models using COVID'19 *Corpus*," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 9, no. 4, pp. 5756-5761, 2020.
- [13] U. Naseem, I. Razzak dan P. W. Eklund, "A survey of pre-processing techniques to improve short-text quality: a case study on hate speech detection on twitter," vol. 80, no. November 2021, pp. 35239-35266, 2020.
- [14] W. Ariannor, E. A. Kusuma, F. Fadilah dan M. Arsyad, "Analyzing User Sentiments in Motor Vehicle Tax Applications Using the Naïve Bayes Algorithm," *Progresif: Jurnal Ilmiah Komputer*, vol. 20, no. 1, pp. 91-101, 2024.
- [15] R. Mohemad, N. N. M. Muhait, N. M. M. Noor dan N. F. A. Mamat, "A Comparative Study of *Stemming* Techniques on the Malay Text," *International Journal of Advanced Computer Science and Applications(IJACSA)*, vol. 14, no. 12, pp. 133-139, 2023.
- [16] X. Yang, K. Yang, T. Cui, M. Chen dan L. He, "A Study of Text *Vectorization* Method Combining Topic Model and Transfer Learning," *Processes*, vol. 10, no. 2, pp. 1-16, 2022.
- [17] D. Z. T. Kannitha, M. Mustafid dan 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.
- [18] S. Mifrah dan E. H. Benlahmar, "Topic Modeling *Coherence*: A Comparative Study between LDA and NMF Models using COVID'19 *Corpus*," *International Journal of Advanced Trends in Computer Science and Engineering (IJATCSE)*, vol. 9, no. 4, pp. 5756-5761, 2020.