

**USULAN KELOMPOK PENELITIAN DOSEN UNPAR
BERDASARKAN PENGELOMPOKAN TOPIK
BIMBINGAN SKRIPSI DENGAN METODE *TOPIC*
*MODELING***

SKRIPSI

Diajukan untuk memenuhi salah satu syarat guna mencapai gelar
Sarjana dalam bidang ilmu Teknik Industri

Disusun oleh :

Nama : Laurensius Avellino
NPM : 6132001054



**PROGRAM STUDI SARJANA TEKNIK INDUSTRI
JURUSAN TEKNIK INDUSTRI
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS KATOLIK PARAHYANGAN
BANDUNG
2024**

**USULAN KELOMPOK PENELITIAN DOSEN UNPAR
BERDASARKAN PENGELOMPOKAN TOPIK
BIMBINGAN SKRIPSI DENGAN METODE *TOPIC
MODELING***

SKRIPSI

Diajukan untuk memenuhi salah satu syarat guna mencapai gelar
Sarjana dalam bidang ilmu Teknik Industri

Disusun oleh:

Nama : Laurensius Avellino

NPM : 6132001054



**PROGRAM STUDI SARJANA TEKNIK INDUSTRI
JURUSAN TEKNIK INDUSTRI
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS KATOLIK PARAHYANGAN
BANDUNG
2024**

**FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS KATOLIK PARAHYANGAN
BANDUNG**



Nama : Laurensius Avellino
NPM : 6132001054
Jurusan : Teknik Industri
Judul Skripsi : USULAN KELOMPOK PENELITIAN DOSEN UNPAR
BERDASARKAN PENGELOMPOKAN TOPIK BIMBINGAN SKRIPSI
DENGAN METODE *TOPIC MODELING*

TANDA PERSETUJUAN SKRIPSI

Bandung, Februari 2024

Ketua Program Studi Sarjana

Teknik Industri

(Dr. Ceicalia Tesavrita, S.P., M.P.T.)

Pembimbing Tunggal

(Dedy Suryadi, S.T., M.S., Ph.D.)



Jurusan Teknik Industri
Fakultas Teknologi Industri
Universitas Katolik Parahyangan

Pernyataan Tidak Mencontek atau Melakukan Tindakan Plagiat

Saya, yang bertanda tangan di bawah ini,

Nama : Laurensius Avellino

NPM : 6132001054

dengan ini menyatakan bahwa skripsi dengan judul :

**“USULAN KELOMPOK PENELITIAN DOSEN UNPAR BERDASARKAN
PENGELOMPOKAN TOPIK BIMBINGAN SKRIPSI DENGAN METODE
TOPIC MODELING”**

adalah hasil pekerjaan saya dan seluruh ide, pendapat atau materi dari sumber lain telah dikutip dengan cara penulisan referensi yang sesuai.

Pernyataan ini saya buat dengan sebenar-benarnya dan jika pernyataan ini tidak sesuai dengan kenyataan, maka saya bersedia menanggung sanksi yang akan dikenakan kepada saya.

Bandung, 12 Januari 2024

Laurensius Avellino
6132001054

ABSTRAK

Lembaga Penelitian dan Pengabdian kepada Masyarakat (LPPM) Universitas Katolik Parahyangan merupakan salah satu Lembaga yang berada dibawah naungan Universitas Katolik Parahyangan yang bertanggung jawab dalam mengembangkan penelitian dan pengabdian masyarakat di lingkungan universitas .LPPM perlu mengetahui kelompok dosen yang memiliki kemampuan yang beririsan dan memiliki keahlian di bidang ilmu yang sama namun lintas jurusan untuk memenuhi salah satu skema penelitian di UNPAR dan memenuhi kebutuhan rencana peleitian kedepannya. Pendekatan dilakukan dengan mengelompokkan skripsi dengan metode *Latent Dirichlet Allocation* (LDA). LDA adalah model probabilistik generatif dari suatu korpus atau kumpulan dokumen. Model generatif dalam LDA ini menciptakan distribusi probabilitas yang mudah dipahami setiap subjek untuk menemukan tema tersembunyi di dalam kumpulan teks. Penilaian dari hasil LDA adalah *coherence score*. Hasil dari LDA adalah kelompok dokumen yang terbagi menjadi 14 topik dengan *coherence score* 0,613780403. Berdasarkan topik yang sudah dikelompokan, akan membentuk kelompok berdasarkan kemampuan dosen.

ABSTRACT

The Research and Community Service Institute (LPPM) of Parahyangan Catholic University is one of the institutes under the auspices of Parahyangan Catholic University that is responsible for developing research and community service within the university. LPPM needs to identify groups of lecturers who have overlapping abilities and expertise in the same field of science but across different departments to fulfill one of the research schemes at UNPAR and meet the needs of future research plans. The approach is to cluster theses using the Latent Dirichlet Allocation (LDA) method. LDA is a generative probabilistic model of a corpus or collection of documents. The generative model in LDA creates a probability distribution that is easy for each subject to understand in order to find hidden themes within the text corpus. The evaluation of the LDA results is the coherence score. The result of LDA is a group of documents divided into 14 topics with a coherence score of 0.613780403. Based on the clustered topics, groups will be formed based on the lecturers' abilities.

KATA PENGANTAR

Puji syukur penulis panjatkan kepada Tuhan Yang Maha Esa atas berkat, rahmat, dan karunia-Nya, penulis dapat menyelesaikan skripsi yang berjudul “Usulan Kelompok Penelitian Dosen UNPAR Berdasarkan Pengelompokan Topik Bimbingan Skripsi Dengan Metode Topic Modeling”. Penelitian dan penulisan dari skripsi dilakukan guna memenuhi syarat untuk mencapai gelar sarjana pada bidang jurusan teknik industri Universitas Katolik Parahyangan. Penulis memperoleh banyak masukan, saran, kritik, dan dukungan dari berbagai pihak dalam proses penyusunan penelitian skripsi ini. Oleh karena itu, penulis ingin mengucapkan terima kasih kepada:

1. Bapak Dedy Suryadi, S.T., M.S., Ph.D. selaku dosen pembimbing yang telah memberikan dukungan, bimbingan, saran, dan masukan selama proses pembuatan laporan skripsi.
2. Orang tua dan keluarga penulis yang selalu mendukung dalam proses kuliah dari awal hingga penulisan skripsi ini.
3. Dr. Yogi Yusuf Wibisono, S.T., M.T. dan Ibu Cynthia Prithadevi Juwono, Ir., M.S. selaku dosen penguji proposal skripsi yang telah memberikan masukan kepada penulis.
4. Dwiky Valeryan Satriyo, Janice Veronica, Imelda Laurensia dan Antonino Veritas selaku teman dekat penulis yang memberi dukungan moral.
5. Aldian Emerson, Keane Bobby, Filbert Jonathan, Fellix Lieguna, Freddy Lorensius, Daniel, Anastasia Vanessa, Jaya Chandra, Marcelyn Angela, Syafrina, Angelica Lintang, Chrissantya Citrani Schalim, Karlo Adibrata dan Michael Richard selaku teman dekat penulis yang memberi dukungan dalam penulisan skripsi.
6. Teman-teman kelas D Teknik Industri 2020 yang sudah berdinamika selama perkuliahan.

Penulis menyadari bahwa masih terdapat kekurangan dalam penelitian ini, oleh sebab itu penulis terbuka kepada semua pihak untuk kritik, saran, dan masukan sehingga dapat membangun serta memperbaiki skripsi ini. Semoga skripsi ini dapat memberikan manfaat bagi pihak yang membaca

Bandung, Januari 2024

Penulis

DAFTAR ISI

KATA PENGANTAR	iii
DAFTAR ISI	v
DAFTAR TABEL	vii
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xv
BAB I PENDAHULUAN	I-1
I.1 Latar Belakang Masalah	I-1
I.2 Identifikasi dan Rumusan Masalah.....	I-3
I.3 Pembatasan Masalah dan Asumsi Penelitian	I-9
I.4 Tujuan Penelitian	I-9
I.5 Manfaat Penelitian	I-10
I.6 Metodologi Penelitian.....	I-10
I.7 Sistematika Penulisan.....	I-13
BAB II TINJAUAN PUSTAKA.....	II-1
II.1 <i>Machine Learning</i>	II-1
II.2 <i>Text Mining</i>	II-2
II.3 <i>Topic Modeling</i>	II-3
II.4 <i>Latent Dirichlet Allocation (LDA)</i>	II-3
II.5 <i>Coherence Score</i>	II-7
BAB III PENGUMPULAN DAN PENGOLAHAN DATA	III-1
III.1 Pengumpulan Data Abstrak	III-1
III.2 <i>Pre-processing Data</i>	III-3
III.2.1 Tokenisasi	III-4
III.2.2 <i>Stopwords</i>	III-6
III.2.3 <i>Lemmatization</i>	III-8
III.3 Penerapan Metode <i>Latent Dirichlet Allocation (LDA)</i>	III-8
III.4 Kombinasi Parameter <i>No Above</i> dan <i>No Below</i>	III-13
III.5 <i>Bigram</i>	III-14
III.6 Hasil Topik LDA	III-16
III.7 Inferensi LDA	III-22

III.8 Usulan Kelompok Penelitian Dosen	III-24
BAB IV ANALISIS	IV-1
IV.1 Analisis Model LDA	IV-1
IV.2 Analisis Nilai <i>No Above</i> dan <i>No Below</i>	IV-4
IV.3 Analisis Bigram	IV-5
IV.4 Analisis Hasil Topik	IV-6
IV.5 Analisis Usulan Kelompok Dosen Pembimbing	IV-7
BAB V KESIMPULAN DAN SARAN	V-1
V.1 Kesimpulan	V-1
V.2 Saran	V-1
DAFTAR PUSTAKA	
LAMPIRAN	
RIWAYAT HIDUP PENULIS	

DAFTAR TABEL

Tabel III.1 Rekapitulasi Data Abstrak.....	III-1
Tabel III.2 Contoh Abstrak	III-2
Tabel III.3 Contoh Proses Tokenisasi	III-4
Tabel III.4 Contoh Daftar <i>Stopwords</i> nltk	III-6
Tabel III.5 Contoh Proses Pembersihan Data.....	III-7
Tabel III.6 <i>Bags-of-Words</i>	III-10
Tabel III.7 Rekapitulasi Nilai <i>Coherence Score</i>	III-13
Tabel III.8 Rekapitulasi <i>Coherence Score</i> Berdasarkan Kombinasi <i>No Above</i> dan <i>No Below</i>	III-14
Tabel III.9 Contoh <i>Bigram</i>	III-15
Tabel III.10 Topik dan Kata-Kata per Topik.....	III-17
Tabel III.11 Judul dan Abstrak Skripsi Topik ke-0 (Sistem Informasi).....	III-20
Tabel III.12 Peluang Dokumen Masuk ke Setiap Topik	III-23
Tabel III.13 Rata-Rata Probabilitas Topik per Pembimbing.....	III-25
Tabel III.14 Reranking Dosen per Topik.....	III-26

DAFTAR GAMBAR

Gambar I.1 Metodologi Penelitian	I-11
Gambar II.1 Ilustrasi <i>Latent Dirichlet Allocation</i>	II-4
Gambar II.2 Representasi Model Grafis dari Model LDA yang Diperhalus	II-6
Gambar III.1 Program <i>Pre-processing Data</i>	III-4
Gambar III.2 Tahap Awal LDA	III-9
Gambar III.3 Model LDA Standar.....	III-11
Gambar III.4 Grafik Jumlah Topik Berdasarkan <i>Coherence Score</i>	III-12
Gambar III.5 Program <i>Bigram</i>	III-16
Gambar III.6 Program Menampilkan Topik.....	III-16
Gambar III.7 Program Inferensi LDA.....	III-22
Gambar III.8 Program Rata-Rata Probabilitas Setiap Pembimbing	III-25
Gambar III.9 Program Peringkat Dosen Pembimbing	III-26

DAFTAR LAMPIRAN

LAMPIRAN A DAFTAR ABSTRAK

LAMPIRAN B DAFTAR *STOPWORDS*

LAMPIRAN C NILAI DOSEN PER TOPIK

LAMPIRAN D PROGRAM PENGOLAHAN DATA

LAMPIRAN E TOP 10 JUDUL DAN ABSTRAK PER TOPIK

BAB I

PENDAHULUAN

Pendahuluan pada penelitian adalah bagian awal dari suatu karya tulis ilmiah yang berisi tentang latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan. Pendahuluan berfungsi untuk memberikan gambaran umum tentang penelitian yang dilakukan. Berikut merupakan latar belakang masalah dari penelitian.

I.1 Latar Belakang Masalah

Pendidikan adalah tindakan yang sengaja direncanakan untuk menciptakan lingkungan pembelajaran dan proses pengajaran sehingga siswa dapat secara aktif mengembangkan kemampuan mereka dalam aspek-aspek seperti spiritualitas, pengendalian diri, kepribadian, kecerdasan, moralitas yang baik, dan keterampilan yang diperlukan untuk kebaikan diri mereka sendiri masyarakat, negara, dan bangsa (Undang-Undang Republik Indonesia Nomor 20 Tahun 2003, Pasal 1).

Universitas Katolik Parahyangan (UNPAR) merupakan salah satu Pendidikan tinggi di Indonesia yang terletak di Bandung. Universitas ini didirikan pada tahun 1955 oleh para uskup di Jawa Barat. UNPAR memiliki semboyan “Bakuning Hyang Mrih Guna Santyaya Bhakti” yang berarti “Berdasarkan Ketuhanan Menuntut Ilmu untuk Dibaktikan kepada Masyarakat”. Universitas Katolik Parahyangan memiliki visi mengembangkan potensi lokal menjadi unggulan pada tataran internasional demi peningkatan martabat manusia dan keutuhan alam ciptaan.

Universitas Katolik Parahyangan memiliki tekad kuat untuk menggali serta mengembangkan potensi lokal yang ada di wilayah Jawa Barat, Indonesia, melalui kegiatan penelitian dan pengabdian kepada masyarakat. Dengan memanfaatkan ilmu, teknologi, dan seni yang dimiliki, universitas ini bertujuan untuk secara kreatif dan kritis memajukan kekuatan, keunggulan, dan unsur keunikan asli (*indigenous*) yang terdapat di wilayah tersebut, sehingga mampu bersaing di tingkat internasional. Selain itu, kegiatan ini juga bertujuan untuk memberikan solusi nyata

terhadap berbagai masalah yang dihadapi oleh masyarakat Indonesia, khususnya di wilayah Jawa Barat.

Untuk menyokong tujuan tersebut, Universitas Katolik Parahyangan memiliki Lembaga Penelitian & Pengabdian Masyarakat (LPPM). Masyarakat (LPPM) UNPAR merupakan lembaga yang bertanggung jawab dalam mengembangkan penelitian dan pengabdian masyarakat di lingkungan Universitas Katolik Parahyangan. LPPM UNPAR memiliki tugas dan fungsi dalam mengembangkan penelitian dan pengabdian masyarakat di lingkungan universitas, serta bertanggung jawab dalam mengelola dana penelitian dan pengabdian masyarakat. Selain itu, LPPM UNPAR juga memiliki beberapa jurnal ilmiah yang diterbitkan. Dalam menjalankan tugasnya, LPPM UNPAR berupaya untuk mengembangkan penelitian dan pengabdian masyarakat yang bermanfaat bagi masyarakat luas. Salah satu tugas utama dari LPPM adalah memfasilitasi penelitian yang dilakukan oleh dosen di Universitas Katolik Parahyangan baik secara pendanaan (internal atau eksternal), serta pemrosesan administrasi seperti membuat kontrak. Terdapat sejumlah skema yang ditawarkan oleh LPPM UNPAR kepada tenaga pengajar unpar untuk dilaksanakan. Terdapat skema yang memiliki syarat dimana penelitian harus dilakukan oleh dosen lintas jurusan.

Seorang dosen tentunya perlu menulis karya ilmiah. Penulisan karya ilmiah oleh dosen tentunya memiliki berbagai manfaat bagi dosen itu sendiri seperti memenuhi tanggung jawab Tridharma, membantu memangku jabatan fungsional, sarana untuk mengembangkan jenjang karir dan memperoleh penghasilan yang lebih baik. Manfaat untuk universitas dan program studi sendiri adalah meningkatkan akreditasi bagi universitas dan program studi berdasarkan jumlah publikasi dan jumlah pemangku jabatan fungsional dan guru besar.

Salah satu faktor yang dapat meningkatkan akreditasi adalah karya ilmiah yang ditulis oleh dosen yang mengajar di universitas tersebut. Karya ilmiah yang ditulis oleh dosen dapat memberikan manfaat bagi masyarakat luas yaitu untuk menyelesaikan masalah, melakukan inovasi dan sebagai sumber pengetahuan.

Penelitian yang dilakukan dosen harus disesuaikan dengan bidang keilmuan dan keahlian dari dosen yang bersangkutan. Seorang dosen pun harus melakukan penelitian setiap tahunnya untuk meningkatkan atau mempertahankan mutu pendidikan Universitas Katolik Parahyangan secara berkala. Proses penelitian tersebut dapat dilakukan secara tunggal atau berkelompok. Dengan

melakukan penelitian berkelompok tentunya proses penelitian akan lebih mudah. Dengan berbagai jenis jurusan yang ada pada Universitas Katolik Parahyangan terdapat peluang dimana dosen yang berasal dari jurusan yang berbeda beririsan dalam hal bidang keilmuan, kemampuan dan kompetensi. LPPM ingin agar dosen yang ada pada Universitas Katolik Parahyangan melakukan penelitian dengan dosen lintas jurusan agar dapat menambah sudut pandang dalam penelitian dan meningkatkan hasil dari penelitian. Dengan hasil penelitian yang baik maka hal tersebut akan menguntungkan dari sisi peneliti atau dosen maupun Universitas Katolik Parahyangan.

I.2 Identifikasi dan Rumusan Masalah

Salah satu cara untuk mempertahankan dan meningkatkan akreditasi dari kampus dan jurusan adalah melalui penelitian yang dilakukan oleh dosen pada Universitas Katolik Parahyangan. LPPM merupakan Lembaga yang memfasilitasi dosen Universitas Katolik Parahyangan. Ibu Maria Fransiska Hanna, S.T., M.T. yang menjabat sebagai Kepala Divisi Lembaga Penelitian dan Pengabdian Masyarakat ingin mengetahui kelompok dosen yang memiliki kemampuan yang beririsan.

Tujuan dari pengelompokan dosen adalah untuk mengetahui apakah terdapat dosen yang memiliki keahlian di bidang ilmu yang sama namun lintas jurusan, sebagai contoh penelitian yang dilakukan oleh dosen lintas jurusan adalah penelitian dengan judul “PENINGKATAN KAPABILITAS DALAM AKUMULASI PENGETAHUAN DAN BERJERARING PADA MAHASISWA DALAM PENCIPTAAN IDE USAHA” (Tahun 2022; ketua: Ir. Catharina Badra Nawangpalupi, Ph.D. (FTI); anggota: Dr. Ceicalia Tesavrita, S.T., M.T. (FTI); Romy Loice, S.T., M.T. (FTI); Dr. Margaretha Banowati Talim, Dra., M.Si (FISIP); Nina Septina, S.P., M.M., M.Phil.(D3)). Dengan adanya usulan nama dosen, LPPM dapat mengajukan nama dosen ketika akan melakukan kerjasama dengan universitas lain. Hal tersebut perlu diketahui sebagai usulan nama dosen dalam melakukan penelitian terutama dalam melaksanakan skema penelitian unggulan perguruan tinggi dan skema penelitian pusat studi yang memiliki jumlah anggota yang terdiri dari 2-5 orang dan minimal berasal dari 2 jurusan. Kedua skema tersebut memiliki tujuan untuk meningkatkan penelitian yang sifatnya lintas disiplin ilmu, sehingga harapannya penelitian dapat dilakukan dengan menimbang

beberapa sudut pandang keilmuan. Perbedaan penelitian unggulan perguruan tinggi adalah jumlah pendanaan yang lebih tinggi dibandingkan dengan golongan penelitian lainnya. Masalah tersebut didapat melalui proses wawancara.

Proses wawancara dilakukan pada hari Jumat, 15 September 2023. Ibu Maria Fransiska Hanna, S.T., M.T memiliki keprihatinan terhadap penelitian yang dilakukan oleh dosen Universitas Katolik Parahyangan dimana hasil penelitian berpotensi memberikan hasil yang lebih baik jika dilakukan oleh dosen lintas jurusan. Ibu Maria Fransiska Hanna, S.T., M.T. menyadari adanya peluang dimana terdapat dosen yang memiliki keahlian di bidang yang sama namun berada pada jurusan yang berbeda di Universitas Katolik Parahyangan. Ibu Maria Fransiska Hanna, S.T., M.T juga menginginkan agar dosen memenuhi tanggung jawab tridharma-nya dengan melakukan penelitian setiap tahunnya. Hal tersebut dapat meningkatkan hasil penelitian dimana terdapat berbagai sudut pandang pada bidang keilmuan yang berbeda dalam suatu penelitian sehingga dapat menghasilkan penelitian yang lebih baik.

LPPM saat ini sedang melakukan pembuatan *roadmap*. Berdasarkan wawancara, *roadmap* dibuat dengan tujuan untuk merencanakan penelitian yang akan dilakukan untuk lima sampai enam tahun kedepan. *Roadmap* sendiri diharapkan akan menghasilkan sesuatu yang bermanfaat. Dalam proses Menyusun dan menjalankan *roadmap*, terdapat kriteria dimana penelitian harus dilaksanakan secara multidisiplin. Pembentukan tim penyusun *roadmap* pun akan dibentuk dengan dosen multidisiplin dengan harapan bahwa berbagai bidang keilmuan yang dimiliki tim *roadmap* dapat memberi sudut pandang yang berbeda ketika melakukan penelitian.

Universitas mendapatkan keuntungan dalam mempertahankan atau meningkatkan akreditasi dari jurusan maupun kampus. Bagi pihak peneliti, keuntungan yang didapat adalah melaksanakan tanggung jawab tridharma serta pengembangan keilmuan dimana ilmu pengetahuan terus berkembang dan dosen sebagai tenaga pengajar harus bisa mengikuti perkembangan ilmu. Dengan ditentukannya kelompok dosen LPPM dapat mempertemukan dosen yang memiliki bidang keilmuan yang beririsan sehingga terdapat potensi kerja sama dalam penelitian. Hal tersebut juga dapat meningkatkan hasil penelitian setiap tahunnya dengan bertambahnya jumlah mitra dosen yang dapat melakukan

penelitian bersama. Saat ini masih terdapat sedikit sekali penelitian yang dilakukan oleh dosen lintas jurusan.

Penelitian yang cukup sering dilakukan di dunia perkuliahan adalah penelitian dosen dan terdapat penelitian yang tidak hanya dilakukan oleh dosen namun dilakukan juga oleh mahasiswa beserta dosen. Bentuk penelitian yang sering dilakukan adalah skripsi untuk program sarjana, thesis untuk pascasarjana dan disertasi untuk gelar doktor. Penelitian-penelitian tersebut tentunya akan ada setiap tahun karena merupakan salah satu syarat wajib yang harus diselesaikan untuk mendapatkan gelar. Dengan jumlah mahasiswa yang banyak, Universitas Katolik Parahyangan tentunya dapat menghasilkan cukup banyak penelitian yang ditulis oleh mahasiswa setiap tahunnya.

Salah satu bagian yang ada pada penelitian adalah abstrak. Abstrak skripsi merupakan rangkuman singkat yang padat, berisi intisari dari seluruh penelitian. Dengan panjang sekitar 150-250 kata, abstrak dirancang untuk memperkenalkan penelitian secara efisien dan menarik minat pembaca. Fungsinya melibatkan memberikan gambaran singkat tentang topik penelitian, menjelaskan pertanyaan penelitian, dan merinci metode yang digunakan, termasuk jenis penelitian, desain, dan instrumen yang dipakai. Abstrak juga menyajikan hasil penelitian secara singkat, menyoroti temuan utama tanpa memasukkan detail mendalam. Abstrak memberikan gambaran keseluruhan yang merangsang minat pembaca untuk lebih mendalami penelitian tersebut. Gambaran yang diberikan oleh abstrak tentunya akan menunjukkan topik dari penelitian dan kata-kata yang berkaitan erat dengan topik dari penelitian. Proses identifikasi dosen dengan bidang keilmuan yang berisikan ditentukan dengan topic modeling. Topic modeling akan menggunakan dokumen sebagai data mentah, kemudian dokumen tersebut akan dikelompokkan berdasarkan topik. Data mentah awal yang digunakan adalah penelitian dosen UNPAR namun LPPM tidak memiliki data abstrak secara terpisah sehingga pendekatan yang dilakukan adalah menggunakan abstrak skripsi yang ada di UNPAR. Akan dilakukan pengelompokan topik skripsi menggunakan abstrak. Setelah pengelompokan dilakukan, setiap skripsi akan diambil dosen pembimbingnya. Nama dosen pembimbing tersebut akan menjadi usulan mitra penelitian. Dengan demikian hal tersebut akan mendapatkan usulan mitra penelitian dosen sesuai dengan kebutuhan dari LPPM UNPAR. Berdasarkan pengamatan awal, terdapat beberapa Skripsi yang memiliki topik yang mirip

namun berasal dari mahasiswa yang berbeda jurusan. Berikut merupakan contoh dari topik Skripsi yang mirip namun berasal dari jurusan yang berbeda.

Tabel I.1 Contoh Topik Skripsi Topik Serupa

No	Judul	Fakultas	Abstrak
1	Pengaruh persepsi konsumen atas content marketing melalui Instagram terhadap minat beli (membership atau paket berlangganan olahraga) konsumen Sosi Sport Club	Fakultas Ekonomi	<p>Penggunaan social media untuk kegiatan digital marketing sangat populer saat ini. Ketatnya rivalitas mewajibkan para pengusaha untuk senantiasa menciptakan bermacam inovasi setiap hari nya sehingga minat beli konsumen bisa didapatkan secara maksimal yang satu di antaranya melalui peningkatan perancangan content marketing. SOSI Sport Club pun menggunakan social media Instagram sebagai salah satu sarana guna memikat serta mempererat komunikasi dengan konsumen. Tujuan dari penelitian ini yakni untuk mendapatkan pengetahuan terkait dengan pengaruh persepsi konsumen atas Content Marketing (reader cognition, sharing motivation, persuasion, decision making, life factors) melalui Instagram SOSI Sport Club terhadap minat beli (membership/paket berlangganan olahraga). Metode penelitian yang digunakan dalam penelitian ini adalah metode eksplanatori. Teknik yang digunakan untuk mengambil sampel dalam penelitian ini adalah teknik non probabily sampling, dengan menggunakan metode purposive sampling. Dalam penelitian ini, jumlah responden yang menjadi sampel adalah 133 responden. Hasil penelitian ini didapatkan Reader cognition, Sharing motivation, Persuasion, Decision Making dan Life Factors berpengaruh secara simultan/memiliki pengaruh secara bersama-sama terhadap minat beli sebesar 91.4%, dan 8.6% dipengaruhi oleh bermacam faktor lain di luar penelitian ini. Agar minat beli konsumen bisa ditingkatkan melalui content Instagram, maka SOSI Sport Club bisa meningkatkan pembuatan konten-kontennya seperti pembuatan konten yang berisikan pengalaman atau testimoni dari seorang influencer/celebgram, suatu list kegiatan workout per minggu, live Instagram atau minivlog, serta konten transformasi yang konsisten update tiap minggu. sehingga konsumen maupun calon konsumen terus mendapatkan manfaat dari informasi yang didapat dari konten instagram SOSI Sport Club, dengan harapan bahwa minat beli konsumen akan terus meningkat. (Kirana, 2021)</p>

(lanjut)

Tabel I.1 Contoh Topik Skripsi Topik Serupa (Lanjutan)

No	Judul	Fakultas	Abstrak
2	Usulan peningkatan minat beli konsumen pada toko alat tulis X	Fakultas Teknologi Industri	Toko X merupakan sebuah toko yang menjual berbagai macam kebutuhan alat tulis kantor. Meskipun sudah lama berdiri, minat beli konsumen pada Toko X cenderung lebih rendah dibandingkan dengan pesaingnya. Dilihat dari jumlah kunjungan yang terjadi, rata-rata jumlah kunjungan setiap harinya pada Toko X lebih sedikit dibandingkan toko lain yang menjual produk sejenis, serta jumlah rata-rata pengunjung yang melakukan pembelian tidak mengalami peningkatan dalam beberapa periode. Sehingga perlu dilakukan tindak lanjut agar Toko X dapat mengetahui faktor apa saja yang perlu dilakukan perbaikan dalam meningkatkan minat beli konsumen. Dalam penelitian ini dilakukan pengujian untuk mengetahui faktor apa saja yang mempengaruhi minat beli pada toko yang menjual kebutuhan alat tulis kantor. Pengujian dilakukan untuk mengetahui pengaruh faktor harga, kualitas layanan, lokasi, dan produk terhadap minat beli. Pengambilan data dilakukan dengan menggunakan kuesioner menghasilkan total 258 responden. Hasil data dari kuesioner tersebut diolah menggunakan metode PLS-SEM (Partial Least Square Modelling). Hasil pengolahan data menunjukkan bahwa faktor harga, kualitas layanan, dan produk memberikan pengaruh positif terhadap minat beli. Selain itu, dari besar efek, hasil observasi, dan hasil wawancara terhadap responden dan pihak Toko X dirancang delapan usulan perbaikan. Hasil usulan tersebut kemudian didiskusikan dan menghasilkan keputusan akhir yaitu 5 usulan diterima meliputi harga, promosi menggunakan media sosial, meningkatkan pemahaman dan pengetahuan pegawai, meningkatkan variasi produk, dan pegawai mengenakan pakaian yang seragam (Suntana,2021).

Tabel I.2 merupakan salah satu contoh skripsi yang memiliki topik yang sama yaitu algoritma genetik namun skripsi pertama merupakan hasil dari mahasiswa fakultas teknologi industri UNPAR dan skripsi kedua merupakan skripsi dari mahasiswa fakultas teknologi informasi dan sains UNPAR.

Dengan jumlah hasil penelitian yang banyak, tentu akan sulit dalam melakukan proses pengelompokan berdasarkan topik sehingga proses

pengelompokan topik akan dilakukan dengan metode *Latent Dirichlet Allocation*. LDA menurut Blei, et al (2003) adalah model probabilistik generatif dari suatu korpus atau kumpulan dokumen. Manfaat utama *Latent Dirichlet Allocation* (LDA) adalah kemampuan interpretasinya yang baik dalam analisis teks. Model generatif dalam LDA ini menciptakan distribusi probabilitas yang mudah dipahami setiap subjek untuk menemukan tema tersembunyi di dalam kumpulan teks. Hal ini memungkinkan pemahaman yang lebih baik tentang tema-tema utama kumpulan dokumen. LDA juga merupakan model probabilistik, yang memungkinkan analisis statistik lebih menyeluruh dan penilaian ketidakpastian dalam alokasi topik. Lebih mudah beradaptasi karena jumlah topik juga dapat ditentukan secara otomatis.

Latent Dirichlet Allocation (LDA) dipilih sebagai metode penelitian yang tepat dengan pertimbangan utama terletak pada kemampuan interpretasi yang baik, model generatif probabilistik, dan kemampuan adaptasi yang tinggi. LDA dapat menghasilkan topik yang mudah dipahami dan diinterpretasikan, memungkinkan identifikasi bidang keilmuan yang beririsan dari dosen-dosen UNPAR. Sebagai model generatif probabilistik, LDA memberikan analisis statistik menyeluruh dan penilaian ketidakpastian, menghasilkan hasil yang lebih akurat dan terpercaya. Kemampuan adaptasi LDA yang tinggi memungkinkan peneliti untuk mengeksplorasi berbagai jumlah topik, mencari solusi optimal, dan menyesuaikan dengan kebutuhan penelitian. Kelebihan lainnya meliputi efisiensi dalam pemrosesan data besar, sederhana dalam implementasi, dan kemudahan interpretasi. Semua ini membuat LDA cocok untuk analisis abstrak skripsi, menyajikan hasil penelitian dengan jelas dan efisien. Berdasarkan latar belakang dan proses identifikasi masalah, dapat dirumuskan masalah sebagai berikut.

1. Bagaimana rancangan model *topic modeling* terbaik yang efektif dalam mengelompokkan topik Skripsi?
2. Bagaimana usulan penelitian dosen UNPAR berdasarkan kemiripan bidang penelitian yang tercermin dari Skripsi yang dibimbingnya?

I.3 Pembatasan Masalah dan Asumsi Penelitian

Setelah identifikasi masalah dan rumusan masalah ditentukan perlu ditentukan Batasan dan asumsi dari penelitian. Batasan dan asumsi dibutuhkan agar penelitian yang dilakukan memiliki fokus dan tujuan yang lebih jelas dan terbatas pada lingkup tertentu. Batasan yang digunakan pada penelitian ini mencakup.

1. Penelitian hanya dilakukan sampai dengan tahap pemberian usulan kepada LPPM Universitas Katolik Parahyangan yang berupa nama tenaga pengajar Universitas Katolik Parahyangan yang sudah dikelompokkan berdasarkan topik skripsi yang mereka bimbing.
2. Data mentah yang digunakan adalah abstrak dari seluruh skripsi, tesis dan disertasi yang ada di Universitas Katolik Parahyangan yang diberikan oleh LPPM Universitas Katolik Parahyangan.
3. Hasil penelitian mahasiswa yang digunakan hanya terbatas dari tahun 2016 sampai dengan 2022.
4. Dokumen penelitian menggunakan Bahasa Indonesia.

Selain Batasan, diperlukan juga asumsi penelitian. Asumsi digunakan untuk menentukan cakupan penelitian. Terdapat pula asumsi yang digunakan untuk menunjang penelitian ini. Asumsi yang digunakan adalah:

1. Dosen yang membimbing setiap skripsi merupakan dosen yang kompeten dan ahli dalam topik penelitian tersebut.
2. Usulan mitra penelitian dosen UNPAR berdasarkan kemiripan bidang penelitian yang tercermin dari skripsi yang dibimbingnya.

I.4 Tujuan Penelitian

Berdasarkan latar belakang dan rumusan masalah yang telah dijelaskan terdapat beberapa tujuan penelitian yang ingin dicapai. Tujuan penelitian ditentukan berdasarkan rumusan dan identifikasi masalah sebelumnya. Berikut merupakan tujuan dari penelitian yang akan dilakukan.

1. Merancang kelompok skripsi berdasarkan topik melalui abstrak dengan metode LDA.
2. Memberikan usulan kelompok penelitian dosen UNPAR berdasarkan kemiripan bidang penelitian yang tercermin dari skripsi yang dibimbingnya.

I.5 Manfaat Penelitian

Pada bagian ini akan dijabarkan mengenai manfaat dari penelitian. Penelitian yang dilakukan tentunya akan memberi manfaat bagi LPPM Universitas Katolik Parahyangan. Manfaat diperoleh dari hasil dan proses penelitian yang dilakukan. Manfaat diharapkan bermanfaat kepada tiga pihak yang terdiri dari LPPM Universitas Katolik Parahyangan, pembaca dan peneliti. Berikut merupakan manfaat yang diperoleh LPPM Universitas Katolik Parahyangan selaku pemilik permasalahan.

1. LPPM Universitas Katolik Parahyangan dapat mengetahui usulan kelompok tenaga pengajar berdasarkan topik skripsi yang dibimbing.
2. LPPM Universitas Katolik Parahyangan mengetahui topik apa saja yang dapat dijadikan bidang penelitian oleh tenaga pengajar lintas jurusan di Universitas Katolik Parahyangan.

Terdapat pula manfaat yang didapat bagi pembaca dan peneliti. Manfaat yang didapat adalah sebagai berikut:

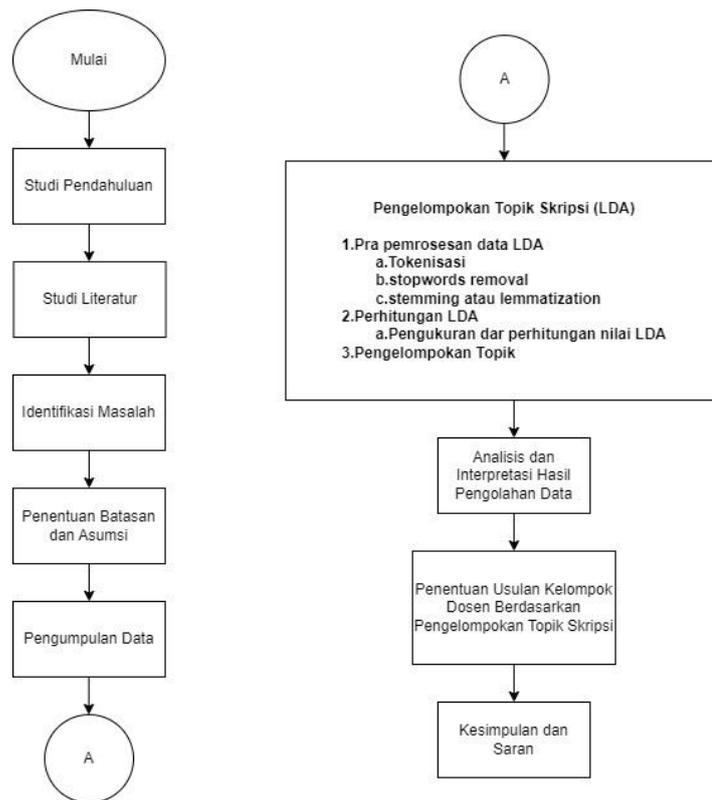
1. Peneliti dapat menerapkan ilmu Teknik industri mengenai *text mining* berupa *topic modeling* dan metode LDA untuk melakukan pengelompokan topik dari kumpulan dokumen.
2. Pembaca dapat memahami permasalahan dan proses penyelesaian pada penelitian ini serta mendapatkan pengetahuan mengenai *text mining*.

I.6 Metodologi Penelitian

Pada bab ini akan dijelaskan mengenai metodologi penelitian. Metodologi penelitian akan memberikan gambaran mengenai hal yang akan dilakukan dalam penelitian serta langkah-langkah penelitian dari mulai, pengolahan data sampai akhir dari penelitian. Metodologi penelitian dapat dilihat pada Gambar I.1. Metodologi penelitian akan membuat penelitian menjadi lebih terstruktur. Metodologi penelitian adalah cara ilmiah untuk mendapatkan data dengan tujuan dan kegunaan tertentu. Metodologi penelitian ini disesuaikan dengan bidang yang akan diteliti dan pada penelitian kali ini *Machine learning* adalah bidang dari penelitian.

Metodologi penelitian terdiri dari Studi Pendahuluan, Studi Literatur, Identifikasi Masalah, Identifikasi Masalah, Penentuan Batasan dan Asumsi,

Pengumpulan Data, Pengelompokan Topik Skripsi menggunakan Metode Latent Dirichlet Allocation, Analisis dan Interpretasi Hasil Pengolahan Data, Penentuan Usulan Kelompok Dosen berdasarkan Pengelompokan Topik Skripsi, dan Kesimpulan dan Saran. Berikut merupakan gambar untuk metodologi penelitian.



Gambar I.1 Metodologi Penelitian

Berikut merupakan penjelasan dari setiap langkah metodologi penelitian penelitian.

1. Studi Pendahuluan

Pada studi pendahuluan akan dilakukan proses pembelajaran terhadap skripsi terutama pada bagian abstrak dikarenakan abstrak akan memberikan gambaran lebih baik terhadap dokumen dibandingkan dengan judul. Proses pembelajaran tersebut dilakukan untuk menentukan latar belakang masalah

2. Studi Literatur

Studi literatur dilakukan dengan mempelajari text mining dengan metode *topic modeling*. Dari *topic modeling* akan digunakan metode *Latent*

Dirichlet Allocation untuk melakukan pengelompokan topik. Akan dilakukan pula pembelajaran mengenai *web scrapping* untuk melakukan proses pengambilan data.

3. Identifikasi dan Perumusan Masalah

Pada tahap ini, akan dijelaskan mengenai identifikasi masalah. Akan dijelaskan mengenai proses identifikasi dan akar permasalahan pada penelitian ini. Berdasarkan identifikasi masalah, akan dibuat rumusan masalah.

4. Penentuan Batasan dan Asumsi

Batasan dan asumsi penelitian perlu ditentukan. Penentuan Batasan bertujuan untuk menentukan area dari penelitian serta memberikan arah jalannya penelitian. Asumsi ditentukan untuk memperlancar proses penelitian dan mengeliminasi kesalahan pada data. Batasan yang digunakan adalah skripsi mahasiswa Universitas Katolik Parahyangan dengan periode 2017 sampai dengan 2023. Asumsi yang digunakan adalah Dosen pembimbing dari setiap skripsi merupakan dosen yang ahli dalam topik tersebut.

5. Pengumpulan Data

Pengumpulan data merupakan proses dimana data mentah dikumpulkan dan akan diolah. Proses pengumpulan data pada penelitian ini akan dilakukan dengan *web scraping*. Data mentah yang digunakan adalah abstrak dari seluruh skripsi Universitas Katolik Parahyangan dari tahun 2016 sampai dengan 2022.

6. Pengelompokan Topik Skripsi (LDA)

Pengelompokan topik skripsi dengan LDA merupakan proses pengolahan data. Pada bagian ini terdapat dua buah proses yaitu Pra pemrosesan data LDA dan ekstraksi topik. Pada pra pemrosesan data, akan dilakukan tokenisasi dimana dokumen akan dipecah menjadi kata atau frasa. Kemudian akan dilakukan *stopwords removal* dimana kata kata yang tidak memiliki makna dan tanda baca akan dihilangkan. Selanjutnya akan dilakukan *stemming* atau *lemmatization* dimana seluruh kata akan ditransformasi menjadi bentuk dasar. Setelah pra pemrosesan data selesai, akan dilakukan perhitungan LDA. Setelah melakukan perhitungan topik akan dikelompokkan dan diambil dosen pembimbing dari

setiap topik sehingga akan didapat nama dosen pembimbing dari setiap topik.

7. Analisis dan Interpretasi Hasil Pengolahan Data

Analisis dan interpretasi hasil akan dilakukan berdasarkan hasil perhitungan LDA dan pengelompokan topik. Hasil yang dianalisis merupakan perhitungan LDA yang akan diinterpretasi menjadi kelompok topik dan kelompok dosen pembimbing.

8. Penentuan Usulan Kelompok Dosen Berdasarkan Pengelompokan Topik Skripsi

Bagian ini akan memberikan usulan mengenai nama nama dosen yang kompeten untuk setiap topik dari berbagai jurusan dan fakultas. Nama dosen tersebut dapat dipasangkan untuk menjadi usulan kerja sama dalam melakukan penelitian dengan topik tertentu.

9. Kesimpulan dan Saran

Kesimpulan akan meringkas hasil penelitian dan diambil berdasarkan analisis dan interpretasi hasil. Kesimpulan tersebut akan menjawab tujuan dari penelitian dan saran akan diberikan dikarenakan penelitian yang dilakukan belum sempurna.

I.7 Sistematika Penulisan

Sistematika penulisan merupakan urutan dalam penulisan laporan penelitian. Sistematika penulisan perlu ditentukan agar penulisan laporan penelitian lebih terstruktur sehingga memiliki alur yang jelas dan dapat menjawab rumusan masalah. Sistematika penulisan penelitian ini akan terdiri dari lima bab. Berikut merupakan penjabaran dari sistematika penulisan laporan penelitian ini.

BAB I PENDAHULUAN

Pendahuluan merupakan bab yang memberi penjelasan awal terhadap penelitian yang dilakukan. Bab pendahuluan terdiri dari latar belakang masalah, identifikasi dan rumusan masalah, pembatasan masalah dan asumsi penelitian, tujuan penelitian, manfaat penelitian, metodologi penelitian dan sistematika penulisan yang akan menjelaskan penelitian yang dilakukan berupa memberi usulan kelompok dosen pembimbing berdasarkan pengelompokan topik skripsi yang dibimbing.

BAB II TINJAUAN PUSTAKA

Tinjauan pustaka akan menjelaskan mengenai teori-teori yang digunakan dalam penelitian yang dilakukan. Teori penelitian terdiri dari *machine learning*, *text mining*, *topic modelling*, *Latent Dirichlet Allocation*, dan *coherence score*.

BAB III PENGUMPULAN DAN PENGOLAHAN DATA

Pengumpulan dan pengolahan data akan berisi proses pengolahan data untuk menjawab tujuan dan rumusan masalah. Data didapat dari LPPM UNPAR dan data abstrak yang diberikan masih terdapat beberapa yang kosong sehingga perlu dilakukan pengisian secara manual melalui repository UNPAR. Pengolahan data menggunakan program dengan bahasa python dan dengan google colab.

BAB IV ANALISIS

Bab analisis akan menjelaskan mengenai analisis dari hasil pengolahan data. Analisis dilakukan dengan melihat nilai coherence dan memberi penilaian secara subjektif dari peneliti. Analisis akan membahas output dari dari program.

BAB V KESIMPULAN DAN SARAN

Bab ini akan menjelaskan mengenai kesimpulan dan saran berdasarkan hasil pengolahan data. Kesimpulan akan disusun untuk menjawab setiap rumusan masalah yang ditentukan untuk LPPM UNPAR.