

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

Skripsi atau tugas akhir menjadi suatu syarat untuk memperoleh gelar Sarjana Strata-1 di perguruan tinggi di Indonesia, tugas akhir merupakan hasil pengamatan terhadap suatu masalah yang terjadi dengan menggunakan metode tertentu di bidang ilmu tersebut (Mahmud,2019), skripsi biasa disebut untuk mahasiswa yang akan mendapatkan gelar S-1 sedangkan untuk bidang Diploma biasa disebut dengan tugas akhir.

Sebelum mencapai tahap skripsi mahasiswa harus menyelesaikan studi minimal 140 sks untuk S-1 dan 110 sks untuk diploma, tahap selanjutnya mahasiswa mengajukan judul penelitian dan Dosen pembimbing untuk proses penyusunan skripsi. Judul skripsi yang diambil oleh mahasiswa biasanya didasarkan dengan mengkaji tugas akhir yang sudah ada dan minat dari mata kuliah yang telah diampu sebelumnya (Panda, 2018). Hal ini akan menentukan tepat atau tidaknya mahasiswa tersebut lulus kuliah, karena judul skripsi yang diambil harus sesuai dengan jenis tema dan program studi yang mahasiswa ambil.

Pendidikan Teknik Informatika dan Komputer (PTIK) merupakan salah satu prodi S-1 di Fakultas Teknik Universitas Negeri Jakarta yang berdiri sejak 31 juli 2009 dengan SK Penyelenggaraan Program Studi: 13300/D/T/K-N/2012 dan pertama kali dibuka pada periode 2010/2011. Pada Program Studi ini mempunyai tiga konsentrasi jurusan di antaranya adalah Rekayasa Perangkat Lunak, Teknik Komputer Jaringan, dan Multimedia.

Pada Prodi Pendidikan Teknik Informatika dan Komputer masing-masing konsentrasi jurusan mempunyai perbedaan terutama dalam hal mata kuliah wajib yang diampu, misalnya konsentrasi Rekayasa Perangkat Lunak mata kuliah wajib yang diampu lebih banyak berhubungan dengan manajemen dalam mengembangkan perangkat lunak dan juga *data mining* disertai dengan beberapa *coding* untuk mengembangkan suatu sistem.

Begitupun konsentrasi Teknik Komputer dan Jaringan di dominasi dengan mata kuliah yang berhubungan dengan jaringan dan alat yang terintegrasi dalam sistem jaringan tertentu, untuk konsentrasi Multimedia lebih banyak berhubungan dengan design, animatografi, editing, untuk memproduksi suatu karya baik itu berupa video, audio, dan gambar dengan menggunakan media digital.

Penulis mencari sumber data dengan membaca dokumen skripsi sebelumnya di *repository* admin Prodi Pendidikan Teknik Informatika Dan Komputer dan *website repository.unj.ac.id*. Pada kenyataannya dalam menyelesaikan studi S-1 mahasiswa Prodi Pendidikan Teknik Informatika dan Komputer tidak semua tema skripsi yang dibuat sesuai dengan konsentrasi jurusan masing-masing, karena ada beberapa mahasiswa yang lebih tertarik di bidang konsentrasi yang berbeda dalam melakukan penelitian karena ketertarikan dengan mata kuliah yang diampu sebelumnya, hal tersebut tentunya membuat banyaknya variasi tema dalam penelitian yang dilakukan di Prodi Pendidikan Teknik Informatika Dan Komputer Universitas Negeri Jakarta dari awal tahun dibuka sampai sekarang.

Oleh karena itu Prodi Pendidikan Teknik Informatika Dan Komputer Universitas Negeri Jakarta belum diketahui jumlah peta topik penelitian skripsi, mengingat semakin lama banyak dokumen skripsi yang masuk, dan masing-masing skripsi yang dibuat mempunyai label sendiri yang tentunya mempunyai topik berbeda-beda dalam satu dokumen, dan teknik pengelompokan hanya dengan melihat topik skripsi secara manual dari judul dan abstrak serta menghabiskan banyak waktu, permasalahan tersebut dapat diselesaikan dengan *Topic Modeling* menggunakan metode *Latent Dirichlet Allocation (LDA)* sehingga pengelompokan topik dilakukan secara otomatis dan lebih terstruktur sesuai komposisinya masing-masing.

Pada tahun ajaran tertentu di Prodi Pendidikan Teknik Informatika Dan Komputer mempunyai kurikulum yang berbeda sesuai kebijakan dari Koordinator Program Studi di mana mata kuliah yang lebih *trend*, dan

dibutuhkan mahasiswa kedepannya akan di perbarui melihat kondisi dari minat mahasiswa dalam mengampu mata kuliah, disisi lain mahasiswa yang mengambil mata kuliah tertentu belum tentu disukai dan sesuai dengan minat mahasiswa, oleh karena itu Prodi Pendidikan Teknik Informatika Dan Komputer Universitas Negeri Jakarta harus mengetahui *trend* mata kuliah yang diminati mahasiswa dari komposisi tema skripsi yang dilakukan sebelumnya agar dalam penyusunan kurikulum lebih baik kedepannya, serta berguna untuk meningkatkan akreditasi dan kurikulum Prodi Pendidikan Teknik Informatika dan Komputer Universitas Negeri Jakarta.

Kesulitan lain yang sering dialami oleh banyak mahasiswa adalah menentukan tema tugas akhir berdasarkan referensi hasil penelitian sebelumnya (Setijohatmo dkk. 2020). Sebagai contoh pada mahasiswa Prodi Pendidikan Teknik Informatika Dan Komputer sering terjadi kesulitan dalam menentukan tema skripsi, karena kurangnya referensi dari penelitian di Prodi Pendidikan Teknik Informatika Dan Komputer sebelumnya, di mana mahasiswa harus mengunduh di *website repository.unj.ac.id* dan harus mempunyai *account* yang terverifikasi perpustakaan terlebih dahulu. Selain itu terdapat perbedaan pendapat antara substansi dan metodologi yang ingin digunakan oleh mahasiswa dan pembimbing mapun penguji dalam menyelesaikan skripsi (Titin., dkk 2019)

Metode yang digunakan biasanya sesuai dengan penelitian relevan sebelumnya dalam jangka waktu maksimal 5 tahun terakhir agar metode yang digunakan lebih update dan data yang digunakan dapat berupa data primer, sekunder, dan tersier di mana masing-masing data dapat diperoleh secara langsung dilapangan dengan metode wawancara, questioner, dan juga menggunakan referensi ilmiah (Riwayah, 2020.).

*Data mining* merupakan suatu bidang ilmu komputer di mana data yang berjumlah besar dari *database* diolah sesuai dengan pola data dan menghasilkan suatu kesimpulan yang dapat mempermudah dalam menafsirkan data (Hossain dkk.,2019). Penerapan *data mining* banyak dilakukan di berbagai bidang misalnya, telekomunikasi, keuangan, sains, dan masih banyak bidang lainnya yang tentunya *data mining* mempunyai

suatu bidang khusus yang lebih spesifik yaitu *text mining* di mana data dapat diklasifikasikan menjadi pola yang menarik dari data tekstual yang berjumlah besar (Roifa,2018).

*Text mining* dapat menyelesaikan permasalahan diantaranya pemrosesan, pengelompokan, dan menganalisa *unstructured text* dalam jumlah besar dengan menggunakan algoritma tertentu, data pada dokumen yang akan dilakukan *text mining* akan dilatih terlebih dahulu di mana data pada dokumen yang sebelumnya tidak terstruktur diubah menjadi terstruktur dengan melakukan beberapa tahap diantaranya: *case folding*, *tokenizing*, *stopwords*, *lemmatization*, *stemming* (Ririd dkk.,2019). Setelah semua data diolah baru dapat disimpulkan hasil dari data yang diinginkan agar lebih mudah dimengerti.

Dalam *text mining* teknik untuk menemukan korelasi antara data dengan dokumen teks, di mana struktur dokumen yang digunakan akan di ambil intisari untuk mengetahui tema yang digunakan sehingga menjadi sebuah informasi (Jelodar, H dkk. 2019) yaitu dengan melakukan pemodelan topik, algoritma yang biasa digunakan adalah *Latent Dirichlet Allocation* (LDA) di mana LDA mengasumsikan dokumen terdiri dari kata-kata dan dipetakan sesuai topik yang berbeda-beda (Doulaty,2019).

Berdasarkan latar belakang masalah tersebut penulis menggunakan metode *Latent Dirichlet allocation* (LDA) untuk pemodelan topik skripsi Prodi Pendidikan Teknik Informatika dan Komputer Universitas Negeri Jakarta.

## 1.2 Identifikasi Masalah

1. Sulitnya mengetahui pengelompokan topik skripsi yang sesuai dengan konsentrasi jurusan yang diminati di Prodi Teknik Informatika Dan Komputer Universitas Negeri Jakarta.
2. Sulitnya mahasiswa dalam menentukan tema skripsi dari referensi tema skripsi sebelumnya di Prodi Teknik Informatika Dan Komputer.

3. Belum diketahui komposisi topik skripsi di Prodi Teknik Informatika Dan Komputer Universitas Negeri Jakarta.

### 1.3 Pembatasan Masalah

1. Studi kasus yang digunakan berfokus pada dokumen skripsi Prodi Pendidikan Teknik Informatika Dan Komputer Universitas Negeri Jakarta.
2. Penelitian ini difokuskan pada dokumen skripsi Mahasiswa Prodi Pendidikan Teknik Informatika Dan Komputer dengan tahun lulus 2017-2022.
3. Metode yang digunakan dalam *Topic Modeling* menggunakan metode *Latent Dirichlet Allocation (LDA)*.

### 1.4 Perumusan Masalah

Berdasarkan latar belakang, identifikasi masalah, dan batasan masalah di atas, maka perumusan masalah yang dibahas pada penelitian ini adalah: Bagaimana mengetahui komposisi topik penelitian yang sesuai dengan isi penelitian dengan menggunakan metode *Latent Dirichlet allocation (LDA)* pada Program Studi Teknik Informatika Dan Komputer Universitas Negeri Jakarta?

### 1.5 Tujuan Penelitian

Mengimplementasikan pemodelan topik menggunakan metode *Latent Dirichlet allocation (LDA)* pada dokumen skripsi Program Studi Pendidikan Teknik Informatika Dan Komputer Universitas Negeri Jakarta.

### 1.6 Manfaat Penelitian

1. Menambah wawasan bagaimana metode *Latent Dirichlet allocation (LDA)* diimplementasikan ke dalam pemodelan topik skripsi Prodi Teknik Informatika Dan Komputer Universitas Negeri Jakarta.
2. Mendapatkan ilmu dan pengetahuan baru terkait *data mining*, *text mining*, dan metode *Latent Dirichlet allocation (LDA)*.

3. Referensi bagi mahasiswa dan program Studi mengenai pengelompokan tema skripsi di Prodi Teknik Informatika Dan Komputer.



*Mencerdaskan dan  
Memartabatkan Bangsa*