

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kehadiran teknologi informasi dalam bidang pendidikan khususnya teknologi dan ilmu pengetahuan memberikan pandangan baru, baik dalam bentuk pikiran atau ide maupun perilaku dalam kegiatan sehari-hari, termasuk ruang lingkup pendidikan (Saputro & Sari, 2020). Pendidikan adalah salah satu faktor yang sangat krusial untuk kehidupan seseorang. Pendidikan dapat dibagi menjadi dua jenis, yaitu pendidikan informal dan pendidikan formal. Pendidikan formal adalah proses belajar yang diselenggarakan oleh lembaga pendidikan, dimulai dari tingkat dasar, menengah, hingga tingkat pendidikan tinggi. Sedangkan pendidikan informal ialah pendidikan yang didapatkan melalui pengalaman, keluarga, masyarakat sekitar, latihan, lingkungan dan sebagainya (Rachmawati & Fibriyani, 2018).

Pengajaran yang efektif dapat menghasilkan alumni perguruan tinggi yang unggul dan berkualitas. Salah satu *output* dari pendidikan di perguruan tinggi ialah wisudawan. Wisudawan merupakan salah satu elemen yang sebelumnya telah menjadi bagian dari perguruan tinggi. Wisudawan mempunyai peran yang sangat penting dalam terwujudnya eksistensi dan kualitas perguruan tinggi (Ritonga et al., 2021).

Universitas Pendidikan Ganesha adalah salah satu lembaga perguruan tinggi di Indonesia, secara rutin menyelenggarakan acara wisuda untuk merayakan keberhasilan para mahasiswanya. Sebagai bagian dari proses evaluasi dan perbaikan berkelanjutan, setiap tahun Undiksha menyebarkan kuesioner online kepada para wisudawan untuk mengetahui kesan dan pesan terkait dengan pengalaman mereka selama masa studi di Undiksha. Tujuan dari pengisian kuesioner adalah untuk memperoleh umpan balik atau tanggapan dari para wisudawan terkait berbagai aspek pengalaman mereka, mulai dari kualitas pendidikan hingga fasilitas dan layanan yang disediakan oleh universitas (NMussafi & Rohaya, 2022). Namun, meskipun kuesioner tersebut telah menjadi bagian penting dari proses evaluasi, belum semua kesan pesan yang dikumpulkan telah diolah oleh tim unit terkait. Hal ini disebabkan oleh jumlah kesan pesan yang cukup banyak yang diterima setiap tahun, sehingga memerlukan waktu dan upaya yang signifikan untuk menganalisisnya secara manual.

Berdasarkan uraian diatas, dibutuhkan sebuah klasifikasi sentimen melalui pesan kesan wisudawan yang dapat dimanfaatkan untuk mengetahui sentimen wisudawan terhadap kinerja perguruan tinggi untuk mempermudah proses evaluasi dalam perbaikan pada lingkungan Undiksha. Adapun kesan pesan wisudawan tersebut dapat diklasifikasikan menjadi sentimen kelas positif, kelas negatif maupun kelas netral menggunakan *text mining*.

Penelitian ini berfokus dalam menemukan suatu topik yang sering dibahas oleh wisudawan pada setiap sentimen positif, negatif dan netral menggunakan metode *Latent Dirichlet Allocation* dan *Non-Negative Matrix Factorization*. NMF dan LDA digunakan untuk mencari topik dari masing-masing sentimen. Kelebihan

dari metode NMF dan LDA mampu mengidentifikasi topik dengan akurat pada sekumpulan data yang besar (Stevens et al., 2012). Pada penelitian ini akan ada dua tujuan utama yang akan dibahas yaitu pertama, klasifikasi sentimen pada dataset kesan pesan wisudawan/wisudawati undiksha menjadi kelas positif, negatif maupun netral dengan menggunakan metode *Support Vector Machine*, *Bidirectional Encoder Representasion Transformer* dan *Long Short Term Memory*. Selain itu, akan diterapkan SMOTE pada kedua metode tersebut untuk mengatasi ketidakseimbangan kelas dalam dataset. Kedua melakukan *topic modeling* pada masing-masing sentimen positif, negatif dan netral untuk mencari topik-topik yang selalu dibicarakan oleh wisudawan pada ketiga kelas sentimen mengimplementasikan metode *Latent Dirichlet Allocation* dan *Non-Negative Matrix Factorization* serta mengukur performa *topic modeling* berdasarkan *coherence score*.

Penelitian serupa dilakukan oleh Putu (2021), penelitian ini bertujuan untuk mengklasifikasi pendapat-pendapat dari wisatawan yang terbagi dalam dua kategori, yaitu positif dan negatif, serta melakukan *topic modeling* pada masing-masing kategori tersebut. *Topic modeling* dilakukan untuk menentukan topik yang paling umum dibicarakan dalam setiap kategori. Langkah-langkah dalam penelitian ini mencakup pengumpulan data, *preprocessing*, transformasi data, pengklasifikasian menggunakan metode *Naive Bayes* serta menerapkan *Latent Dirichlet Allocation* untuk analisis tema. Pengujian performa model *Naive Bayes* menunjukkan hasil akurasi mencapai 92%, 100% untuk presisi, *recall* sebesar 83,84%, dan spesifisitas mencapai 100%. Sementara itu, hasil analisis tema menggunakan LDA menunjukkan bahwa topik dengan koherensi tertinggi untuk

kelas positif adalah topik 8 dengan nilai *coherence score* mencapai 0,613, sedangkan pada kelas negatif adalah topik 12 dengan nilai *coherence score* sebesar 0,528.

Penelitian selanjutnya dilakukan oleh Alkaff (2021) yang bertujuan untuk mengembangkan sistem yang dapat mengklasifikasikan laporan secara otomatis berdasarkan topiknya untuk layanan pemerintah Indonesia yang dikenal sebagai *Lapor!* Proses ini mengintegrasikan metode *Latent Dirichlet Allocation* dengan *Support Vector Machine*. Hasil uji yang dilakukan dengan metode pembagian data dengan perbandingan 70:30 bahwa model menunjukkan performa yang memuaskan, hasil akurasi mencapai 79,85%, *recall* 72,37%, presisi 79,98%, dan *F1-Score* 74,67%.

Merujuk pada penelitian-penelitian terdahulu, peneliti terdorong untuk melakukan penelitian dengan judul **“Klasifikasi Sentimen dan *Topic Modeling* terhadap Kesan Pesan Wisudawan Undiksha”**. Pada penelitian ini akan menambah beberapa metode klasifikasi seperti metode SVM, LSTM dan BERT. Pada penelitian sebelumnya hanya menggunakan sentimen positif dan sentimen negatif. Maka penelitian ini akan di kembangkan menjadi tiga sentimen yaitu positif, negatif dan netral. Dikarenakan data yang akan digunakan tidak seimbang antara sentimen positif, negatif dan netral. Maka akan dilakukan *resampling* data dengan menggunakan teknik *SMOTE*.

1.2 Rumusan Masalah

Berdasarkan informasi yang telah dijelaskan sebelumnya pada latar belakang, adapun beberapa permasalahan yang dapat diuraikan sebagai berikut.

1. Bagaimana hasil perbandingan metode SVM, LSTM dan BERT dengan menerapkan SMOTE dan Non-SMOTE pada kesan pesan wisudawan Undiksha?
2. Bagaimana hasil *topic modeling* menggunakan metode LDA dan NMF dalam mengidentifikasi topik-topik yang sering dibahas dalam kesan pesan wisudawan undiksha pada masing-masing kelas positif, negatif dan netral?

1.3 Batasan Masalah

Adapun batasan masalah pada penelitian ini yaitu dibatasi dengan cakupan sebagai berikut.

1. Data yang akan diambil merupakan data kesan pesan wisudawan Universitas Pendidikan Ganesha pada tahun 2021 sampai tahun 2023.
2. Metode yang digunakan yaitu metode SVM, LSTM, BERT, LDA dan NMF.
3. Data yang digunakan merupakan data berbahasa Indonesia.
4. Klasifikasi kesan pesan wisudawan menjadi kelas positif, negatif dan netral.

1.4 Tujuan Penelitian

1. Untuk mengetahui bagaimana hasil perbandingan metode SVM, LSTM dan BERT dengan menerapkan SMOTE dan Non-SMOTE pada kesan pesan wisudawan Undiksha?

2. Untuk mengetahui hasil *topic modeling* menggunakan metode LDA dan NMF dalam mengidentifikasi topik-topik yang sering dibahas dalam kesan pesan wisudawan undiksha pada masing-masing kelas positif, negatif dan netral.

1.5 Manfaat Penelitian

a. Manfaat Teoritis

Penulis berharap bahwa penelitian ini dapat meningkatkan pemahaman atau pengetahuan, serta mampu mengimplementasikan ilmu yang didapatkan selama perkuliahan. Penelitian ini juga dapat digunakan sebagai bahan acuan atau referensi untuk peneliti lainnya yang memiliki penelitian yang sejenis dimasa depan.

b. Manfaat Praktis

1. Dapat mengetahui kesan pesan wisudawan selama menempuh pendidikan di Universitas Pendidikan Ganesha.
2. Dapat mengetahui topik-topik yang sering dibahas oleh wisudawan terhadap data pesan dan kesan.
3. Dapat mengetahui hasil penerapan teknik SMOTE pada data yang tidak seimbang.