

BAB I

PENDAHULUAN

Pada bab ini, latar belakang penelitian akan dijelaskan secara mendalam, meliputi permasalahan yang dihadapi serta alasan mengapa solusi yang diajukan sangat penting. Selanjutnya, akan dibahas permasalahan utama yang menjadi fokus penelitian beserta tujuan yang ingin dicapai. Terakhir, akan dijelaskan manfaat penelitian ini baik dalam lingkup akademik maupun industri, serta keterbatasan yang ada dalam penelitian ini.

1.1 Latar Belakang

Kekerasan terhadap perempuan merupakan salah satu bentuk pelanggaran hak asasi manusia yang masih marak terjadi di berbagai negara, termasuk Indonesia. Fenomena ini tidak hanya menyisakan luka fisik dan psikis bagi para korban, tetapi juga mencerminkan ketimpangan sosial dan budaya yang telah mengakar kuat dalam masyarakat. Berdasarkan laporan tahunan Komnas Perempuan tahun 2021, tercatat sebanyak 8.234 kasus kekerasan terhadap perempuan di Indonesia, dengan 79% di antaranya merupakan kekerasan dalam rumah tangga, dan sebagian besar korbannya adalah istri[1]. Angka tersebut kemungkinan jauh lebih rendah dibandingkan jumlah kasus yang sebenarnya terjadi di lapangan, mengingat masih banyak korban yang memilih diam karena takut, malu, atau tidak tahu harus melapor ke mana.

Kekerasan terhadap perempuan dapat terjadi dalam berbagai bentuk, mulai dari kekerasan fisik, psikologis, seksual, hingga kekerasan ekonomi dan sosial. Kasus-kasus tersebut dapat berlangsung di lingkungan domestik seperti rumah tangga, maupun di ruang publik seperti tempat kerja, jalan umum, atau institusi pendidikan. WHO (*World Health Organization*) bahkan memperkirakan bahwa satu dari tiga perempuan di dunia pernah mengalami kekerasan fisik atau seksual, baik oleh pasangan maupun oleh orang lain[2]. Tingginya angka ini memperlihatkan bahwa kekerasan terhadap perempuan bukanlah masalah individu, melainkan persoalan struktural yang berkaitan erat dengan norma gender, relasi kuasa, dan sistem hukum yang belum sepenuhnya berpihak pada korban.

Salah satu kendala terbesar dalam menangani kasus kekerasan terhadap perempuan di Indonesia adalah keterbatasan dalam proses dokumentasi dan analisis laporan. Di tingkat daerah, Dinas Pemberdayaan Perempuan dan Perlindungan Anak (DPPA) menjadi lembaga yang bertanggung jawab menangani kasus-kasus kekerasan yang dilaporkan oleh masyarakat. Laporan yang diterima oleh DPPA biasanya berbentuk narasi kronologis dari korban atau pendampingnya, yang berisi penjabaran tentang kejadian kekerasan, jenis tindakan, pelaku, serta dampak yang dirasakan. Namun, laporan-laporan tersebut sering kali sangat beragam dari segi gaya penulisan, bahasa, dan detail informasi, sehingga menyulitkan petugas untuk melakukan klasifikasi atau pengelompokan secara sistematis. Hal ini menjadi tantangan tersendiri.

Kondisi ini menunjukkan perlunya sebuah pendekatan yang mampu membantu proses ekstraksi informasi dari laporan naratif secara efisien dan akurat. Salah satu metode yang telah banyak digunakan dalam pengolahan dokumen yang tidak terstruktur adalah *Latent Dirichlet Allocation (LDA)*. LDA merupakan model topik berbasis probabilistik yang mampu mengelompokkan kumpulan kata dalam dokumen ke dalam topik-topik tertentu yang memiliki keterkaitan semantik. Dalam konteks laporan kekerasan terhadap perempuan, metode ini dapat digunakan untuk mengidentifikasi tema-tema umum yang sering muncul dalam kronologi korban, seperti kekerasan dalam rumah tangga, pelecehan seksual, atau kekerasan fisik di tempat umum. Dengan demikian, LDA dapat menjadi alat bantu untuk memahami pola-pola kekerasan yang terjadi di masyarakat tanpa harus membaca seluruh laporan satu per satu secara manual[3].

Namun, proses analisis tidak berhenti pada ekstraksi topik semata. Agar sistem dapat memberikan manfaat yang lebih luas, terutama dalam konteks klasifikasi laporan baru secara otomatis, maka diperlukan pendekatan tambahan yang mampu melakukan klasifikasi berbasis informasi topik. Dalam hal ini, metode *Zero-shot classification (ZSC)* menjadi solusi yang relevan. ZSC merupakan pendekatan klasifikasi yang dirancang untuk bekerja secara efektif meskipun hanya memiliki sedikit data pelatihan. Dengan menggunakan representasi topik dari LDA sebagai fitur, sistem ZSC dapat mengklasifikasikan laporan baru ke dalam kategori-kategori kekerasan yang telah dikenali sebelumnya, seperti KDRT, kekerasan seksual, atau

kekerasan ekonomi. Pendekatan ini sangat berguna dalam konteks lembaga seperti DPPA, yang sering kali tidak memiliki data terlabel dalam dan tidak terstruktur untuk keperluan pelatihan model klasifikasi konvensional.

Beberapa penelitian terdahulu menunjukkan bahwa kombinasi LDA dan metode klasifikasi berbasis pembelajaran minimal mampu memberikan hasil yang akurat dalam domain teks sosial. Sebuah studi memanfaatkan LDA untuk menganalisis ratusan ribu dokumen kekerasan terhadap perempuan dan berhasil mengidentifikasi lima topik utama yang berulang dalam narasi[4]. Sementara itu, Conneau mengembangkan XLM-RoBERTa, sebuah model transformer multilingual yang dilatih pada lebih dari 100 bahasa dan dioptimalkan untuk tugas pengenalan inferensi teks lintas bahasa (XNLI), menjadikannya sangat cocok untuk tugas *zero-shot classification* dalam berbagai konteks linguistik. Temuan-temuan ini menguatkan bahwa pendekatan berbasis representasi semantik dan klasifikasi *zero-shot* layak diterapkan dalam konteks laporan kekerasan terhadap perempuan di Indonesia, di mana ketersediaan data berlabel sangat terbatas[5].

Penelitian ini dilakukan dengan menggunakan dataset laporan korban kekerasan dari DP3AKB Kabupaten Jember. Laporan-laporan tersebut berisi informasi tentang identitas korban, waktu kejadian, dan kronologi kekerasan yang dialami. Fokus wilayah penelitian ini dibatasi hanya pada Kabupaten Jember untuk memperdalam analisis secara lokal dan kontekstual. Dengan menggunakan metode LDA untuk ekstraksi topik dan ZSC untuk klasifikasi, sistem yang dikembangkan diharapkan mampu membantu DP3AKB dalam mengelola laporan secara lebih efisien, mengidentifikasi pola kekerasan yang sering terjadi, serta memberikan dasar yang kuat dalam perumusan kebijakan perlindungan perempuan berbasis data. Melalui pendekatan ini, penanganan kasus kekerasan tidak hanya menjadi lebih cepat dan sistematis. Selain itu, hasil penelitian ini akan diimplementasikan ke dalam sebuah website, yang memungkinkan DP3AKB untuk mengakses dan menganalisis data secara *real-time*, memberikan manfaat yang lebih besar dalam upaya peningkatan perlindungan perempuan di Kabupaten Jember.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, rumusan masalah dalam penelitian ini meliputi:

1. Bagaimana penerapan metode *Latent Dirichlet Allocation* (LDA) dalam mengekstraksi topik utama dari laporan kasus kekerasan terhadap perempuan di Kabupaten Jember?
2. Bagaimana penerapan metode *Zero-shot Classification* (ZSC) dalam mengklasifikasikan laporan kekerasan berdasarkan representasi topik dari LDA?
3. Bagaimana performa kombinasi metode LDA dan ZSC dalam klasifikasi topik laporan kekerasan berdasarkan metrik evaluasi seperti akurasi, presisi, *recall*, dan *F1-score*?

1.3 Tujuan Penelitian

Penelitian ini bertujuan untuk menerapkan metode Latent Dirichlet Allocation (LDA) dalam menganalisis laporan kasus kekerasan yang dialami oleh perempuan di Kabupaten Jember dan memanfaatkan hasil representasi topik tersebut sebagai dasar untuk melakukan klasifikasi secara otomatis menggunakan Zero-Shot Classification (ZSC). Penelitian ini mengeksplorasi bagaimana LDA dapat mengidentifikasi topik-topik utama dari laporan korban kekerasan yang diterima oleh Dinas Pemberdayaan Perempuan dan Anak (DPPA), serta bagaimana ZSC dapat mengklasifikasikan laporan-laporan tersebut berdasarkan representasi topik yang dihasilkan. Selain itu, penelitian ini juga bertujuan untuk mengevaluasi performa kombinasi metode LDA dan ZSC melalui metrik evaluasi seperti akurasi, presisi, recall, dan F1-score, serta melakukan pencarian parameter optimal LDA menggunakan pendekatan Bayesian Optimization untuk meningkatkan kualitas hasil topik.

1.4 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat teoritis maupun praktis. Secara lebih spesifik, manfaat yang diharapkan dari penelitian ini mencakup:

1. Bagi Penulis

Manfaat penelitian ini bagi penulis adalah untuk mengaplikasikan ilmu yang telah dipelajari selama semasa studi, khususnya di bidang sistem kecerdasan

rekayasa. Selain itu, manfaat penelitian ini juga untuk menambah wawasan dan pemahaman mengenai penerapan teknik pemodelan topik dan klasifikasi dalam analisis berbasis teks, khususnya dalam konteks kasus-kasus kekerasan yang dialami oleh perempuan. Penelitian ini juga dapat menjadi referensi untuk penelitian lanjutan mengenai penerapan *Latent Dirichlet Allocation* dan *Zero-Shot Classification*.

2. Bagi Masyarakat Umum

Manfaat penelitian ini bagi masyarakat umum adalah untuk meningkatkan kesadaran masyarakat mengenai isu-isu yang dihadapi oleh perempuan saat ini, sehingga dapat membuat masyarakat lebih peduli dan responsif terhadap masalah gender.

3. Bagi Pemerintah

Manfaat penelitian ini bagi pemerintah adalah memberikan wawasan yang lebih mendalam tentang isu-isu yang sering terjadi pada perempuan, khususnya di Kabupaten Jember, sehingga dapat membantu dalam merumuskan kebijakan yang lebih ketat dan efektif untuk melindungi hak-hak perempuan.

4. Bagi Pekerja di Bidang Teknologi

Manfaat penelitian ini bagi pekerja di bidang teknologi adalah untuk menjadi referensi bagi pengembangan aplikasi berbasis teknologi dalam bidang sosial, seperti analisis berita atau layanan publik yang berfokus pada perlindungan perempuan.

1.5 Batasan Masalah

Agar penelitian ini lebih terfokus maka penelitian ini membatasi beberapa aspek sebagai berikut:

1. Penelitian ini hanya berfokus pada kasus kekerasan yang dialami oleh perempuan di Kabupaten Jember yang tercatat dalam laporan yang didapatkan dari korban.
2. Penelitian ini hanya mengklasifikasikan kasus kekerasan terhadap perempuan ke dalam empat jenis kekerasan, yaitu kekerasan fisik, kekerasan seksual, kekerasan psikis, dan kekerasan verbal, berdasarkan laporan yang diterima dari Dinas Pemberdayaan Perempuan, Perlindungan Anak, dan Keluarga Berencana Kabupaten Jember.

3. Data yang digunakan berjumlah 300 laporan kasus kekerasan dari 2018 sampai 2024 yang telah ditranskrip dari buku laporan manual menjadi file digital di *Microsoft Excel*.
4. Evaluasi performa klasifikasi menggunakan *latent dirichlet allocation* dan *zero-shot classification* dibatasi pada *coherence score* dan metrik akurasi, presisi, *recall*, dan *F1-score*.