

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pemanfaatan teknologi informasi terus menunjukkan peningkatan yang cepat pada berbagai ranah termasuk pendidikan, sosial, dan kesehatan. Salah satu sektor yang paling terdampak oleh kemajuan ini adalah sektor kesehatan [1], di mana peningkatan akses informasi serta kesadaran masyarakat terhadap hak atas layanan yang berkualitas mendorong mereka untuk lebih aktif menyuarakan pendapat dan pengalamannya. Hal ini terlihat dari semakin banyaknya ulasan yang dipublikasikan di berbagai platform digital, seperti media sosial, *Google Maps*, dan situs ulasan lainnya. Ulasan tersebut mencerminkan persepsi masyarakat terhadap layanan kesehatan yang diterima dan menjadi sumber informasi penting bagi penyedia layanan dalam melakukan evaluasi dan peningkatan mutu pelayanan.

Sebagai penyelenggara layanan kesehatan utama, rumah sakit memiliki tanggung jawab besar dalam menjaga mutu pelayanannya. Mutu pelayanan yang optimal tidak semata-mata berpengaruh pada kepuasan pasien, melainkan juga mampu memperkuat reputasi rumah sakit di tengah persaingan industri kesehatan yang semakin kompetitif. Oleh karena itu, rumah sakit perlu tanggap terhadap berbagai masukan masyarakat, termasuk ulasan yang disampaikan secara digital, sehingga dapat dilakukan perbaikan yang tepat sasaran dan berkesinambungan [2].

RSUD Syarifah Ambami Rato Ebu (Syamrabu) Bangkalan merupakan rumah sakit rujukan di wilayah Madura yang mengalami peningkatan signifikan jumlah kunjungan, dari 196.873 pada tahun 2023 menjadi 241.803 pada tahun 2024. Lonjakan ini menunjukkan meningkatnya kepercayaan masyarakat, namun juga menghadirkan tantangan dalam menjaga konsistensi mutu pelayanan. Berdasarkan data ulasan publik di *Google Maps*, RSUD Syamrabu Bangkalan memperoleh rating rata-rata 4,7 dari 5 bintang dengan ribuan ulasan dari masyarakat. Meskipun nilai ini menunjukkan tingkat kepuasan yang cukup tinggi, ulasan yang diberikan masyarakat tetap beragam, baik yang bersifat apresiatif maupun yang berisi masukan terhadap kualitas layanan rumah sakit[3].

Variasi ulasan tersebut menggambarkan adanya persepsi yang beragam dari pasien terhadap mutu layanan rumah sakit. Akan tetapi, sebagian besar ulasan

pasien memiliki format tidak teratur, menggunakan ragam bahasa yang berbeda, dan jumlahnya terus meningkat seiring waktu. Kondisi tersebut membuat analisis manual sulit dilakukan karena memerlukan banyak waktu dan sumber daya. Analisis terhadap ulasan publik penting bagi rumah sakit karena dapat membantu memahami persepsi pasien serta menjadi dasar evaluasi dan peningkatan mutu layanan secara berbasis data. Oleh karena itu, diperlukan metode analisis teks yang mampu mengelompokkan dan mengidentifikasi tema utama dalam ulasan secara otomatis dan efisien.[4].

Salah satu metode yang dapat diterapkan ialah *Topic Modeling*. Teknik ini merupakan sebuah cara untuk metode klasterisasi non-hierarkis yang secara otomatis dapat mengidentifikasi dan mengelompokkan dokumen ke dalam topik-topik tersembunyi berdasarkan kemunculan kata[5]. Algoritma yang paling sering dimanfaatkan dalam penerapan *Topic Modeling* adalah *Latent Dirichlet Allocation(LDA)*. *LDA* mampu mengekstraksi topik-topik tersembunyi dari kumpulan dokumen dan memberikan distribusi probabilistik terhadap setiap topik dalam dokumen. Algoritma ini dinilai unggul dalam menangani data teks dalam jumlah besar dan telah digunakan dalam berbagai penelitian analisis teks, termasuk di sektor publik. Berbeda dengan metode *Clustering* seperti *K-Means* yang hanya mengelompokkan teks berdasarkan kesamaan kata, atau *Sentiment Analysis* yang berfokus pada penilaian emosi positif dan negatif, *LDA* mampu mengidentifikasi topik-topik tersembunyi secara otomatis sehingga lebih efektif untuk menemukan tema utama dalam ulasan publik [6].

Beberapa kajian terdahulu yang relevan dan sejumlah penelitian sebelumnya yang berkaitan dengan topik ini antara lain diteliti oleh [7], Riset ini ditujukan guna mengidentifikasi sekaligus menganalisis tren riset di bidang kesehatan di Indonesia dengan memanfaatkan *Topic Modeling Latent Dirichlet Allocation(LDA)*. Hasil penelitian menunjukkan bahwa algoritma *LDA* mampu melakukan Analisis topik yang diterapkan terhadap judul penelitian kesehatan di Indonesia. Dari proses pemodelan tersebut, diperoleh dua kelompok utama, yaitu topik bersifat umum serta topik yang berkaitan dengan penyakit. Visualisasi hasil pemodelan, seperti *Wordcloud* dan *PyLDavis*, memperlihatkan persebaran kata dan frasa yang relevan terhadap judul penelitian, sehingga memudahkan dalam memahami distribusi topik.

Berdasarkan penilaian Partisipan penelitian mencakup dokter atau tenaga medis, peneliti yang bergerak di bidang kesehatan, serta akademisi, sebanyak 94,1% menyatakan bahwa hasil pemodelan sangat baik, sementara 5,9% menilai baik. Topik yang diidentifikasi meliputi faktor-faktor kesehatan, status gizi, tekanan darah, rumah sakit, ibu hamil, pelayanan kesehatan, Puskesmas, ekstrak etanol, kesehatan masyarakat, HIV/AIDS, diabetes, dan demam berdarah. Hasil ini menunjukkan keberhasilan metode *LDA* dalam memodelkan tren penelitian kesehatan di Indonesia yang kemudian bisa dimanfaatkan sebagai landasan untuk pengembangan penelitian lebih lanjut serta sebagai referensi dalam analisis tren penelitian di bidang kesehatan.

Penelitian lainnya oleh [8]. Pada penelitian ini menggunakan pendekatan *LDA* untuk menganalisis ulasan hotel di Kabupaten Bandung. Penelitian ini mengindikasikan bahwa masukan dari ulasan digital dapat dijadikan sumber informasi guna menelaah aspek daya tarik hotel di Kabupaten Bandung. Melalui metode ekstraksi topik dengan *LDA*, tiga faktor utama teridentifikasi, yaitu “Fasilitas dan Lingkungan”, “Aktivitas dan Hiburan”, serta “Kualitas dan Pelayanan”. Analisis varians menunjukkan tidak ada perbedaan signifikan antar faktor tersebut. Pada Grand Sunshine Resort & Convention Soreang Bandung, faktor “Fasilitas dan Lingkungan” menjadi dominan dengan bobot sebesar 30,56%, sedangkan di Sutan Raja Hotel and Convention Centre, faktor “Aktivitas dan Hiburan” sebesar 32,06% dan “Kualitas dan Pelayanan” sebesar 33,22% menjadi faktor utama. Secara keseluruhan, “Kualitas dan Pelayanan” muncul sebagai faktor paling berpengaruh dalam membentuk daya tarik hotel, meskipun perbedaan bobot antar faktor relatif kecil di keempat hotel yang diteliti.

Beberapa penelitian sebelumnya menunjukkan keberhasilan *penerapan LDA*, misalnya untuk menganalisis tren penelitian kesehatan di Indonesia serta mengidentifikasi faktor-faktor utama dalam ulasan hotel. Namun, Sampai saat ini, belum terdapat penelitian yang berfokus secara khusus menganalisis ulasan pasien RSUD Bangkalan menggunakan *LDA*. Hal ini menjadi peluang untuk menggali topik-topik dominan dalam ulasan pasien, sehingga dapat memberikan informasi strategis bagi rumah sakit dalam evaluasi dan peningkatan mutu pelayanan.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk menerapkan *Topic Modeling* menggunakan *algoritma Latent Dirichlet Allocation(LDA)* pada ulasan pasien RSUD Syarifah Ambami Rato Ebu Bangkalan. Hasil penelitian ini dapat mengungkap topik-topik utama yang menjadi perhatian pasien serta menjadi dasar evaluasi bagi pihak manajemen rumah sakit dalam upaya meningkatkan mutu pelayanan kesehatan secara berkelanjutan.

1.2 Rumusan Masalah

Penelitian ini memiliki rumusan masalah yang dapat dijabarkan yang terdiri dari permasalahan, Metode Usulan dan pertanyaan penelitian. :

1.2.1 Permasalahan

Saat ini, masyarakat semakin aktif memberikan ulasan mengenai pelayanan rumah sakit melalui platform digital, seperti media sosial, forum, dan situs ulasan. Data ulasan tersebut bersifat tidak terstruktur dan jumlahnya besar, sehingga sulit dianalisis secara manual. RSUD Syarifah Ambami Rato Ebu Bangkalan, sebagai rumah sakit rujukan daerah, perlu memahami persepsi masyarakat terhadap layanan yang diberikan agar dapat melakukan perbaikan secara tepat sasaran. Untuk itu, diperlukan pendekatan otomatis, yakni *Topic Modeling* yang berbasis *algoritma Latent Dirichlet Allocation(LDA)*, yang dapat mengekstraksi topik utama dari kumpulan ulasan. Penelitian ini berfokus pada penerapan *LDA* dalam mengidentifikasi dan menganalisis topik dominan dari ulasan masyarakat terkait pelayanan rumah sakit.

1.2.2 Solusi Permasalahan

Penelitian ini menggunakan metode *Latent Dirichlet Allocation(LDA)*, yaitu *algoritma Topic Modeling* berbasis probabilistik, untuk menganalisis dan mengelompokkan topik-topik utama yang muncul dari ulasan mengenai pelayanan di RSUD Syarifah Ambami Rato Ebu Bangkalan. Metode *LDA* mampu mengidentifikasi distribusi kata dalam dokumen terhadap topik serta menemukan tema-tema tersembunyi secara otomatis dari data teks ulasan yang tidak terstruktur.

1.2.3 Pertanyaan Penelitian

Berdasarkan permasalahan dan solusi yang diidentifikasi, maka pertanyaan penelitian yang dapat diusulkan dalam penelitian ini adalah sebagai berikut :

1. Topik apa saja yang muncul dan bagaimana distribusi topik pada ulasan pasien terhadap pelayanan RSUD Syarifah Ambami Rato Ebu Bangkalan?
2. Berapa jumlah topik terbaik yang diperoleh berdasarkan nilai *Coherence Score* dari hasil pemodelan *LDA*?
3. Bagaimana pengaruh variasi nilai parameter α (*Alpha*) dan β (*Beta*) terhadap perubahan nilai *Coherence Score* dalam penerapan model *LDA*?

1.3 Tujuan dan Manfaat

1.3.1 Tujuan Penelitian

Berdasarkan permasalahan, solusi, dan pertanyaan penelitian, maka tujuan penelitian ini adalah sebagai berikut:

1. Mengidentifikasi topik-topik utama serta menganalisis distribusi topik pada ulasan pasien terhadap pelayanan di RSUD Syarifah Ambami Rato Ebu Bangkalan.
2. Menentukan jumlah topik terbaik yang dihasilkan berdasarkan nilai coherence pada pemodelan *Latent Dirichlet Allocation(LDA)*.
3. Menganalisis pengaruh variasi parameter *Alpha* dan *Beta* terhadap nilai coherence dalam penerapan model *LDA*.

1.3.2 Manfaat

Adapun manfaat yang diperoleh dari penelitian ini adalah sebagai berikut, baik dari sisi akademis maupun praktis:

1. Akademis: Penelitian ini memberikan kontribusi terhadap pengembangan ilmu pengetahuan di bidang *Natural Language Processing (NLP)*, khususnya dalam penerapan metode *Topic Modeling* menggunakan algoritma *Latent Dirichlet Allocation(LDA)* pada analisis teks di domain kesehatan. Hasil penelitian ini dapat menjadi referensi bagi penelitian selanjutnya yang berfokus pada pemodelan topik, analisis opini publik, serta penerapan *NLP* dalam peningkatan mutu layanan kesehatan.
2. Praktis: Penelitian ini memberikan insight bagi manajemen RSUD Syarifah Ambami Rato Ebu Bangkalan dalam memahami isu dan persepsi yang

paling sering muncul pada ulasan pasien. Temuan topik-topik utama yang dihasilkan dapat dijadikan bahan evaluasi dan dasar perbaikan pelayanan rumah sakit secara berkelanjutan.

1.4 Batasan Masalah

Batasan masalah yang dapat diterapkan agar penelitian ini tidak terlalu luas dan terarah, adalah sebagai berikut:

1. Data ulasan yang digunakan berasal dari satu platform, yaitu *Google Maps*, khusus untuk RSUD Syarifah Ambami Rato Ebu Bangkalan, dengan jumlah 2004 data karena platform ini menyediakan jumlah ulasan yang cukup banyak, terbuka, dan representatif dibandingkan sumber lainnya.
2. Proses *Preprocessing* teks dibatasi pada tahapan *Cleaning*, *Case Folding*, *Tokenization*, *Normalization*, *Stopword Removal*, dan *Stemming* untuk memastikan konsistensi pengolahan data.
3. Analisis topik dilakukan dengan menggunakan algoritma *Latent Dirichlet Allocation(LDA)* tanpa mengombinasikannya dengan metode lain seperti *Sentiment Analysis* atau *Clustering*.
4. Data ulasan yang dianalisis hanya mencakup teks berbahasa Indonesia, tanpa menyertakan bahasa daerah, slang, maupun istilah nonbaku yang sulit diproses secara otomatis.
5. Hasil pemodelan topik divisualisasikan menggunakan *Wordcloud* dan *PyLDAvis* untuk mempermudah interpretasi tema-tema yang terbentuk.

1.5 Sistematika Penulisan

Penyusunan sistematika penelitian ini bertujuan untuk memandu penulis menulis dengan alur yang runtut, menjaga fokus analisis, dan tetap konsisten pada tujuan penelitian. Berikut adalah susunan sistematika laporan penelitian yang digunakan:

BAB 1 PENDAHULUAN

Bab ini merupakan latar belakang penelitian, rumusan masalah, permasalahan yang dihadapi, metode yang diusulkan, pertanyaan penelitian, tujuan dan manfaat penelitian, batasan masalah yang dibahas, serta sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Pada bab ini menjelaskan mengenai kumpulan teori ilmiah yang melandasi masalah penelitian yang terdiri dari *Text Mining*, *Topic Modeling*, Algoritma *Latent Dirichlet Allocation (LDA)* dan penelitian terkait.

BAB III METODE PENELITIAN

Bab ini memaparkan pendekatan penelitian yang dilakukan, mulai dari penggunaan dataset hingga penjelasan alur kerja sistem.

BAB IV HASIL DAN PEMBAHASAN

menyajikan hasil penelitian, mulai dari statistik data awal, hasil pemodelan *LDA*, kata kunci dan distribusi topik, interpretasi topik, hingga analisis perbandingan dengan penelitian terdahulu. Bab ini juga memuat diskusi mengenai insight yang diperoleh dan relevansinya terhadap fokus penelitian.

BAB V PENUTUP

berisi ringkasan temuan penelitian, pencapaian tujuan penelitian, serta rekomendasi untuk penelitian selanjutnya atau penerapan hasil penelitian dalam praktik. Terakhir, Daftar Pustaka memuat semua referensi ilmiah, buku, artikel, dan sumber data yang digunakan, yang pada penelitian ini disusun secara otomatis menggunakan aplikasi Mendeley. Sistematika ini memastikan laporan penelitian tersusun secara terstruktur, mudah dipahami, dan fokus pada analisis yang relevan.

DAFTAR PUSTAKA

Pada bab ini berisi tentang keterangan mengenai sumber rujukan yang digunakan dalam penyusunan tugas akhir. Pada penelitian ini daftar pustaka disusun secara otomatis dengan menggunakan aplikasi Mendeley