

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi telah mendorong transformasi digital pada berbagai sektor pelayanan publik, termasuk sektor ketenagalistrikan. Digitalisasi layanan memungkinkan penyedia layanan meningkatkan aksesibilitas, kecepatan pelayanan, dan penyampaian informasi kepada masyarakat. Salah satu implementasinya adalah PLN Mobile yang dikembangkan oleh PT PLN (Persero) sebagai kanal layanan digital untuk memenuhi kebutuhan pelanggan, mulai dari pembelian token, pembayaran tagihan, perubahan daya, pencatatan meter, hingga penyampaian pengaduan layanan.

(Syafrizal et al., 2024) Selain berfungsi sebagai media transaksi dan pelayanan, PLN Mobile menghasilkan data ulasan pelanggan pada Google Play Store. Ulasan tersebut merekam pengalaman pengguna secara langsung, baik dalam bentuk apresiasi, kritik, maupun keluhan terkait aplikasi dan layanan kelistrikan. Informasi semacam ini relevan untuk mengevaluasi kualitas layanan karena mencerminkan persepsi pengguna terhadap fitur, gangguan, dan layanan yang mereka gunakan (Syafrizal et al., 2024).

Jumlah ulasan yang besar menimbulkan tantangan analitis karena data tersebut berbentuk teks tidak terstruktur dan memiliki variasi bahasa yang tinggi. Pengguna dapat menyampaikan keluhan menggunakan singkatan, istilah tidak baku, kesalahan penulisan, maupun struktur kalimat yang beragam. Akibatnya, informasi penting seperti jenis gangguan, fitur yang bermasalah, aplikasi yang terdampak, dan layanan kelistrikan yang berkaitan sulit diidentifikasi serta dihubungkan secara sistematis apabila hanya dianalisis secara manual.

Penelitian ini menggunakan data sekunder yang berasal dari penelitian Fauzan Fuadi Pasaribu (2025) berjudul "*Implementasi NER Pada Ulasan Pengguna Aplikasi PLN Mobile Menggunakan FastText dan BiLSTM-CRF*". Dataset tersebut bersumber dari ulasan PLN Mobile pada Google Play Store. Pada proses pengumpulan awal

diperoleh 522.000 ulasan, kemudian diseleksi berdasarkan periode Januari 2020 hingga Desember 2024 menjadi 451.612 ulasan. Setelah proses *cleaning* untuk menghapus data kosong, duplikat, ulasan tidak relevan, dan ulasan yang tidak memenuhi batas minimum kata, diperoleh 202.532 ulasan sebagai dataset akhir. Berbeda dengan penelitian terdahulu tersebut yang memanfaatkan data untuk *Named Entity Recognition*, penelitian ini menggunakan dataset yang sama untuk membangun *Domain Knowledge Graph* pengaduan layanan kelistrikan dan melakukan analisis sentimen berbasis konteks.

(Rizkiyanto et al., 2024) *Natural Language Processing (NLP)* dapat digunakan untuk mengolah ulasan tidak terstruktur melalui tahapan pembersihan teks, ekstraksi fitur, ekstraksi informasi, dan klasifikasi. Pada objek PLN Mobile, (Rizkiyanto et al., 2024) menerapkan preprocessing, *TF-IDF*, dan *Random Forest* serta memperoleh *F1-score* 93,14%, sedangkan (Syafrizal et al., 2024) menggunakan *Naive Bayes* dan *K-Nearest Neighbor* dengan akurasi terbaik 77,69%. Temuan tersebut menunjukkan bahwa ulasan PLN Mobile dapat dianalisis secara komputasional, tetapi keluaran penelitian tersebut terutama berupa kelas sentimen dan belum merepresentasikan keterhubungan semantik antar objek pengaduan.

(Asri et al., 2022; Tambunan & Hapsari, 2021) Penelitian lain oleh (Tambunan & Hapsari, 2021) menggunakan *text mining*, *Word Cloud*, *Network Explorer*, dan *VADER* untuk mengidentifikasi pola opini pengguna New PLN Mobile. (Asri et al., 2022) juga memanfaatkan *VADER*, *TF-IDF*, dan *Naive Bayes* untuk pelabelan serta klasifikasi sentimen. Meskipun pendekatan tersebut mampu menggambarkan polaritas dan pola opini, informasi mengenai hubungan antara gangguan aplikasi, fitur aplikasi, aplikasi, dan layanan kelistrikan belum direpresentasikan dalam suatu struktur pengetahuan yang eksplisit.

Keterbatasan tersebut penting karena satu ulasan dapat memuat beberapa konsep yang saling berkaitan. Sebagai contoh, pernyataan "aplikasi error saat login sehingga gagal membeli token" tidak hanya menunjukkan sentimen negatif, tetapi juga memuat gangguan berupa error, fitur login, objek aplikasi, dan layanan token. Klasifikasi sentimen konvensional dapat menentukan polaritas ulasan, tetapi tidak secara langsung

menjelaskan hubungan semantik antarkonsep tersebut. Oleh karena itu, diperlukan representasi yang mampu mempertahankan konteks dan keterhubungan informasi di dalam ulasan.

(Hogan et al., 2022) *Knowledge Graph* digunakan dalam penelitian ini karena mampu merepresentasikan pengetahuan melalui entitas, relasi, dan atribut dalam struktur graf. Entitas direpresentasikan sebagai *node*, sedangkan hubungan antarentitas direpresentasikan sebagai *edge* atau *relationship*, sehingga informasi yang sebelumnya tersebar dalam teks dapat ditelusuri berdasarkan keterkaitan semantiknya. (Hogan et al., 2022) menjelaskan bahwa *Knowledge Graph* menyediakan kerangka representasi yang menghubungkan entitas dan relasi untuk membentuk pengetahuan yang lebih terstruktur.

(Zhong et al., 2023) Agar struktur *Knowledge Graph* tidak terbentuk secara arbitrer, penelitian ini menggunakan *Ontology Modeling* sebagai skema konseptual domain. Ontologi mendefinisikan kelas, hubungan, atribut, dan batasan konsep yang digunakan sebagai acuan dalam proses ekstraksi entitas dan relasi. Dalam penelitian ini, ontologi difokuskan pada empat kelas utama, yaitu Aplikasi, Gangguan_Aplikasi, Fitur_Aplikasi, dan Layanan_Kelistrikan. Pendekatan berbasis ontologi dipilih agar setiap entitas dan relasi yang masuk ke *Knowledge Graph* memiliki makna yang konsisten dengan domain pengaduan PLN Mobile (Hogan et al., 2022).

(Chen et al., 2022; Li & Wang, 2023; Wang et al., 2021) Pembangunan *Knowledge Graph* dari teks tidak terstruktur sejalan dengan tahapan *knowledge acquisition* dan *knowledge refinement* yang dibahas oleh (Zhong et al., 2023). Dalam penelitian ini, ulasan pelanggan diproses melalui *text preprocessing*, *Entity Extraction* berbasis kamus, *Rule-Based Relation Extraction*, pembentukan *RDF Triple*, dan implementasi pada Neo4j. Rangkaian tersebut memungkinkan informasi hasil ekstraksi diubah menjadi struktur *subject-predicate-object* dan selanjutnya direpresentasikan sebagai *node* dan *relationship* dalam basis data graf.

Pada domain kelistrikan, *Knowledge Graph* telah digunakan untuk integrasi data, pemeliharaan, diagnosis gangguan, dan peningkatan interpretabilitas informasi (Wang et al., 2021; Chen et al., 2022; Li & Wang, 2023). Namun, penelitian tersebut lebih

banyak berfokus pada data teknis sistem tenaga, data simulasi, atau data operasional. Sementara itu, kajian terhadap PLN Mobile lebih dominan menggunakan analisis sentimen dan opini. Kondisi ini menunjukkan adanya celah penelitian pada pemanfaatan ulasan pelanggan nyata sebagai sumber pengetahuan untuk membangun *Knowledge Graph* pengaduan layanan kelistrikan.

Berdasarkan celah tersebut, penelitian ini membangun Domain *Knowledge Graph* pengaduan layanan kelistrikan dari 202.532 ulasan pelanggan PLN Mobile sebagai *Knowledge Base*. Kontribusi penelitian tidak berhenti pada pembentukan graf, tetapi juga mengevaluasi kualitas *Entity Extraction* dan *Relation Extraction* menggunakan *precision*, *recall*, dan *F1-score* untuk memastikan bahwa struktur pengetahuan yang dibentuk memiliki tingkat ketepatan dan ketercakupan yang terukur.

Knowledge Base yang dihasilkan selanjutnya dimanfaatkan untuk memberikan konteks pada analisis sentimen. Sentimen dilabelkan berdasarkan rating pengguna menjadi negatif, netral, dan positif, kemudian teks direpresentasikan menggunakan *TF-IDF* dan diklasifikasikan dengan *Multinomial Naive Bayes* serta *Random Forest*. Hasil sentimen kemudian dikaitkan dengan kelas ontologi melalui kamus entitas yang konsisten dengan tahap *Entity Extraction*. Dengan mekanisme tersebut, analisis tidak hanya menunjukkan polaritas ulasan, tetapi juga menunjukkan konteks objek atau kategori pengaduan yang berkaitan dengan sentimen tersebut.

Dengan demikian, penelitian ini mengintegrasikan pengolahan teks, *Ontology Modeling*, *Knowledge Graph*, dan analisis sentimen dalam satu alur penelitian. Pendekatan ini diharapkan menghasilkan representasi pengaduan pelanggan yang lebih terstruktur dan interpretatif, sehingga hubungan antara gangguan aplikasi, fitur aplikasi, aplikasi, dan layanan kelistrikan dapat ditelusuri melalui *Knowledge Graph* serta digunakan sebagai konteks dalam membaca kecenderungan sentimen pelanggan PLN Mobile.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijabarkan diatas, dapat dirumuskan permasalahan sebagai berikut:

- 1) Bagaimana merancang *Ontology* domain pengaduan layanan kelistrikan berdasarkan data ulasan pelanggan PLN *Mobile*?
- 2) Bagaimana membangun dan mengevaluasi *Knowledge Graph* pengaduan layanan kelistrikan melalui tahapan *text preprocessing*, *Ontology Modeling*, *Entity Extraction*, *Relation Extraction*, *RDF Triple Construction*, dan implementasi menggunakan *Neo4j*, berdasarkan struktur entitas, relasi, *RDF Triple*, *node*, dan *relationship* yang dihasilkan?
- 3) Bagaimana menganalisis sentimen ulasan pelanggan PLN *Mobile* dengan memanfaatkan *Knowledge Graph* yang telah dibangun sebagai *Knowledge Base* untuk memberikan konteks terhadap hasil analisis sentimen?

1.3 Tujuan Penelitian

Berdasarkan latar belakang yang telah dijabarkan di atas, adapun tujuan penelitian ini sebagai berikut:

- 1) Merancang *Ontology* domain pengaduan layanan kelistrikan berdasarkan data ulasan pelanggan PLN *Mobile*.
- 2) Membangun dan mengevaluasi *Knowledge Graph* pengaduan layanan kelistrikan melalui tahapan *text preprocessing*, *Ontology Modeling*, *Entity Extraction*, *Relation Extraction*, *RDF Triple Construction*, dan implementasi menggunakan *Neo4j*.
- 3) Melakukan analisis sentimen ulasan pelanggan PLN *Mobile* dengan memanfaatkan *Knowledge Graph* sebagai *Knowledge Base*.

1.4 Manfaat Penelitian

Berdasarkan latar belakang yang telah dijabarkan di atas, adapun manfaat penelitian ini sebagai berikut:

1.4.1 Manfaat Akademis

- 1) Penelitian ini dapat memberikan kontribusi terhadap pengembangan ilmu di bidang *Natural Language Processing* dan *Knowledge Graph*, khususnya dalam pengolahan data ulasan pelanggan berbahasa Indonesia pada domain layanan kelistrikan.

- 2) Penelitian ini dapat menjadi referensi bagi penelitian selanjutnya yang ingin membangun *Domain Knowledge Graph* dari data teks tidak terstruktur serta memanfaatkannya sebagai *Knowledge Base* untuk mendukung analisis sentimen.
- 3) Penelitian ini dapat memperkaya kajian mengenai pemanfaatan *Ontology Modeling*, *Entity Extraction*, *Relation Extraction*, *RDF Triple*, dan analisis sentimen dalam merepresentasikan serta menganalisis informasi pengaduan pelanggan.

1.4.2 Manfaat Praktis

- 1) Penelitian ini dapat membantu penyedia layanan kelistrikan dalam memahami informasi pengaduan pelanggan secara lebih terstruktur melalui hubungan antara pelanggan, gangguan aplikasi, fitur aplikasi, dan layanan kelistrikan.
- 2) *Knowledge Graph* yang dibangun dapat digunakan sebagai *Knowledge Base* untuk mengidentifikasi hubungan antarentitas dan jenis gangguan atau keluhan yang terdapat pada ulasan pelanggan PLN *Mobile*.
- 3) Hasil analisis sentimen dapat memberikan informasi mengenai kecenderungan sentimen positif, netral, dan negatif pada ulasan pelanggan PLN *Mobile* serta konteks pengaduan yang berkaitan dengan sentimen tersebut.

1.4.3 Manfaat Sosial

- 1) Penelitian ini diharapkan dapat mendukung peningkatan kualitas layanan kelistrikan digital melalui pemanfaatan ulasan pelanggan sebagai sumber informasi pengaduan.
- 2) Penelitian ini dapat membantu mendorong pelayanan yang lebih *responsif* terhadap keluhan masyarakat, khususnya terkait gangguan aplikasi, token, tagihan, pembayaran, kWh, meteran, dan layanan kelistrikan lainnya.
- 3) Dengan adanya representasi pengaduan dan analisis sentimen yang lebih terstruktur, penelitian ini diharapkan dapat mendukung peningkatan kualitas pelayanan dan komunikasi antara penyedia layanan kelistrikan dengan masyarakat.

1.5 Ruang Lingkup Masalah

Ruang lingkup masalah yang akan dibahas dalam penulisan, yaitu sebagai berikut:

- 1) Data yang digunakan dalam penelitian ini adalah dataset ulasan pelanggan PLN *Mobile* dari Google Play Store yang telah tersedia dalam bentuk file terstruktur.
- 2) Penelitian berfokus pada pembangunan dan evaluasi *Domain Knowledge Graph* pengaduan layanan kelistrikan serta pemanfaatannya sebagai *Knowledge Base* untuk mendukung analisis sentimen ulasan pelanggan PLN *Mobile*.
- 3) Tahapan pembangunan *Knowledge Graph* meliputi *text preprocessing*, *Ontology Modeling*, *Entity Extraction*, *Relation Extraction*, *RDF Triple Construction*, pembangunan *Knowledge Graph* menggunakan *Neo4j*, dan evaluasi.
- 4) *Text preprocessing* yang digunakan meliputi *cleaning*, *case folding*, normalisasi kata, *tokenization*, *stopword removal*, dan *stemming*.
- 5) *Ontology* yang dibangun difokuskan pada domain pengaduan layanan kelistrikan pada aplikasi PLN *Mobile*, dengan *class* utama yaitu Aplikasi, Gangguan_Aplikasi, Fitur_Aplikasi, dan Layanan_Kelistrikan.
- 6) *Ontology* yang dibangun difokuskan pada domain pengaduan layanan kelistrikan pada aplikasi PLN *Mobile*, dengan *class* utama yaitu Aplikasi, Gangguan_Aplikasi, Fitur_Aplikasi, dan Layanan_Kelistrikan.
- 7) Hasil *Knowledge Graph* direpresentasikan dalam bentuk *node* dan *relationship* menggunakan *Neo4j* serta dievaluasi berdasarkan hasil ekstraksi entitas, relasi, dan struktur graf yang dihasilkan.
- 8) Analisis sentimen dibatasi menjadi tiga kelas sentimen, yaitu negatif, netral, dan positif, dengan pelabelan berdasarkan *rating* pengguna PLN *Mobile*.
- 9) Analisis sentimen menggunakan representasi fitur TF-IDF dengan membandingkan algoritma Multinomial Naive Bayes dan Random Forest.
- 10) Evaluasi model analisis sentimen dilakukan menggunakan *accuracy*, *precision*, *recall*, *F1-score*, dan *confusion matrix*.

- 11) Pemanfaatan *Knowledge Graph* sebagai *Knowledge Base* dalam analisis sentimen dilakukan dengan mengaitkan hasil sentimen dengan *class* pada *Ontology*, yaitu Aplikasi, Gangguan_Aplikasi, Fitur_Aplikasi, dan Layanan_Kelistrikan.

1.6 Sistematika Penulisan

Sistematika penulisan dalam penelitian ini disusun sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, ruang lingkup penelitian, dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Bab ini membahas penelitian terdahulu dan landasan teori yang berkaitan dengan ulasan pelanggan, *Natural Language Processing*, *text preprocessing*, *Knowledge Graph*, *Ontology Modeling*, *Entity Extraction*, *Relation Extraction*, *RDF Triple*, Neo4j, dan evaluasi *Knowledge Graph*, *Knowledge Graph* sebagai Dasar Analisis Sentimen, Dasar Pemilihan Pendekatan Penelitian, Dasar Pemilihan Pendekatan Penelitian, Alasan Menggunakan *Ontology Modeling*, Alasan Menggunakan *Rule-Based Relation Extraction*, Alasan Menggunakan *RDF Triple*, Alasan Menggunakan Neo4j, Alasan Tidak Menggunakan *Large Language Model*, Alasan Tidak Menggunakan *Knowledge Graph Embedding*, Alasan Menggunakan *Dictionary-Based Entity Extraction*, Alasan Tidak Menggunakan *Open Information Extraction*, Alasan Menggunakan *OWL (Web Ontology Language)*, Hubungan Semantik Dalam *Knowledge Graph*, Web Scraping

BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan metode penelitian yang digunakan, yaitu *Ontology -Based Domain Knowledge Graph Construction*. Bab ini juga menjelaskan waktu dan lokasi penelitian, data penelitian, sumber data penelitian, karakteristik dataset, metode penelitian, alur penelitian, tahapan *preprocessing*, *Ontology Modeling*, *Entity*

Extraction, Relation Extraction, RDF Triple, Domain Knowledge Graph (Neo4j), dan evaluasi Metode, Metode Analisis data.

BAB IV HASIL DAN PEMBAHASAN

Bab ini membahas hasil implementasi penelitian, mulai dari Hasil penelitian, Hasil *web scraping*, Hasil *teks preprocessing*, *word cloud teks preprocessing*, Hasil *Ontology modeling*, Hasil *Entity extraction*, *Word cloud entity extraction*, Hasil *relation extraction*, Hasil *RDF Triple*, Hasil *Domain Knowledge Graph*, Pembahasan.

BAB V PENUTUP

Bab ini berisi kesimpulan dari hasil penelitian serta saran untuk pengembangan penelitian selanjutnya.