BAB I
PENDAHULUAN

1.1 Latar Belakang

Efisiensi pelayanan administrasi di lingkungan instansi pemerintahan sering kali terhambat oleh keterbatasan sumber daya manusia dalam merespons volume pertanyaan yang fluktuatif dan kompleks. Pada unit kerja Inrenbang, fenomena penumpukan antrean pertanyaan (*bottleneck*) di layanan *helpdesk* menjadi masalah krusial yang mempengaruhi kinerja operasional harian. Staf administrasi kepegawaian sering kali harus menangani pertanyaan yang diajukan oleh pegawai internal secara manual, di mana satu pertanyaan kerap mengandung beberapa intensi sekaligus, seperti permintaan informasi mengenai prosedur cuti yang digabungkan dengan pertanyaan terkait tunjangan kinerja. Kondisi ini menuntut adanya mekanisme respons yang tidak hanya cepat tetapi juga akurat dalam memetakan berbagai topik dalam satu kalimat input. Keterlambatan dalam memberikan respons yang relevan berdampak langsung pada kepuasan pegawai dan efektivitas manajemen waktu di lingkungan kerja, sehingga diperlukan solusi otomasi yang mampu mengakselerasi proses tanya jawab tersebut tanpa menambah beban kerja staf *helpdesk* (Langgeng et al., 2023).

Kompleksitas bahasa alami yang digunakan oleh pegawai dalam mengajukan pertanyaan menuntut sistem klasifikasi yang lebih canggih daripada sekadar pencocokan kata kunci sederhana. Masalah utama yang dihadapi dalam pengembangan sistem otomasi di Inrenbang adalah karakteristik data pertanyaan yang bersifat *multi-label*, di mana satu dokumen teks dapat dikategorikan ke dalam lebih dari satu kelas label secara bersamaan. Pendekatan klasifikasi teks tradisional yang bersifat *single-label* sering kali gagal menangkap nuansa ini, menyebabkan jawaban yang diberikan oleh sistem menjadi tidak lengkap atau tidak relevan. Oleh karena itu, penerapan metode *problem transformation* seperti *Binary Relevance* dan *Classifier Chains* menjadi sangat relevan untuk diuji kemampuannya dalam memecah kompleksitas tersebut. Urgensi penelitian ini terletak pada kebutuhan untuk membangun model yang mampu memahami korelasi antar label dalam konteks administrasi yang spesifik dan tertutup, guna memastikan setiap bagian dari pertanyaan pegawai terjawab dengan tuntas (Nofriani & Kurniawan, 2021a).

Penelitian ini menjadi semakin penting mengingat batasan operasional yang ada pada *chatbot* yang akan dikembangkan, di mana basis pengetahuan (*knowledge base*) yang digunakan terbatas pada *dataset* sampel internal yang disediakan oleh atasan magang, bukan bersumber dari data global atau internet terbuka. Hal ini menciptakan tantangan tersendiri

karena model harus mampu melakukan generalisasi yang baik dari jumlah data yang terbatas dan sangat spesifik terhadap regulasi internal Inrenbang. Jika sistem mengandalkan pencarian global, risiko hallusinasi atau ketidakrelevanan jawaban terhadap kebijakan internal akan sangat tinggi. Oleh sebab itu, fokus penelitian diarahkan pada bagaimana mengoptimalkan algoritma *machine learning* untuk mengekstraksi pola dari *dataset* lokal tersebut agar *chatbot* dapat berfungsi sebagai asisten virtual yang andal dalam ekosistem tertutup. Keberhasilan dalam lingkup ini akan membuktikan bahwa efisiensi tinggi dapat dicapai tanpa memerlukan *big data* berskala masif, melainkan melalui pemilihan metode klasifikasi yang tepat sasaran (Wang, 2023).

Implementasi metode *Classifier Chains* dan *Binary Relevance* dipilih sebagai fokus komparasi karena kedua metode ini menawarkan pendekatan yang berbeda dalam menangani independensi dan ketergantungan label. *Binary Relevance* bekerja dengan memecah masalah *multi-label* menjadi beberapa masalah klasifikasi biner yang independen, yang menawarkan kecepatan komputasi namun sering kali mengabaikan hubungan antar label. Di sisi lain, *Classifier Chains* berusaha memodelkan ketergantungan antar label dengan merantaikan pengklasifikasi, sehingga prediksi label sebelumnya menjadi fitur tambahan bagi label berikutnya. Analisis komparatif antara kedua metode ini pada *dataset* administrasi kepegawaian Inrenbang akan memberikan wawasan baru mengenai *trade-off* antara akurasi prediksi gabungan dan efisiensi waktu pemrosesan. Argumentasi utamanya adalah bahwa dalam dokumen administratif, label sering kali memiliki korelasi kuat (misalnya, 'cuti' sering muncul bersama 'absensi'), sehingga metode yang memperhitungkan korelasi label diprediksi akan memberikan performa yang lebih superior (Arslan & Cruz, 2023).

Data tren global menunjukkan adanya pergeseran signifikan menuju adopsi *Artificial Intelligence* (AI) dalam manajemen sumber daya manusia (SDM) untuk meningkatkan efisiensi operasional. Berdasarkan laporan industri terkini, organisasi yang mengadopsi solusi NLP untuk layanan internal mampu memangkas waktu respons hingga 60% dan mengurangi beban kerja administratif rutin hingga 40%. Namun, tantangan teknis tetap ada, terutama terkait dengan ambiguitas teks dan kelangkaan data berlabel yang berkualitas tinggi dalam domain spesifik seperti pemerintahan daerah atau unit kerja khusus. Masalah yang mendasari penelitian ini bukan hanya pada ketersediaan teknologi, tetapi pada bagaimana teknologi tersebut diadaptasi untuk menangani *multi-intent* pada data berskala kecil (*small data*). Tanpa penanganan yang tepat terhadap aspek *multi-label*, tingkat kegagalan *chatbot* dalam memahami konteks pengguna akan tetap tinggi, yang pada gilirannya akan menurunkan tingkat kepercayaan pengguna terhadap sistem otomasi (Wang, 2024).

Tinjauan terhadap penelitian terdahulu menunjukkan bahwa pengembangan *chatbot* untuk layanan publik telah banyak dilakukan, namun sebagian besar masih berfokus pada klasifikasi *single-label* atau pencarian jawaban berbasis aturan (*rule-based*). Beberapa studi telah berhasil mengembangkan *chatbot* akademik dan layanan masyarakat yang mampu menjawab pertanyaan umum dengan akurasi yang baik. Temuan utama dari riset-riset ini mengindikasikan bahwa penggunaan algoritma seperti Naïve Bayes atau SVM cukup efektif untuk teks pendek. Namun, keterbatasan mendasar dari penelitian-penelitian tersebut adalah ketidakmampuan model untuk menangani input pengguna yang kompleks dan mengandung banyak permintaan sekaligus. Ketika dihadapkan pada skenario *multi-label*, performa model *single-label* menurun drastis karena sistem dipaksa memilih satu kategori dominan saja, mengabaikan informasi penting lainnya yang terkandung dalam pesan pengguna (Cortés-Cediel et al., 2023).

Sementara itu, penelitian lain yang berfokus pada klasifikasi teks *multi-label* telah mengeksplorasi penggunaan *Deep Learning* seperti LSTM atau BERT untuk menangani ketergantungan label yang kompleks. Temuan dari studi-studi ini menunjukkan bahwa model berbasis *Deep Learning* mampu mencapai skor F1 yang sangat tinggi pada *dataset* publik berskala besar seperti Reuters atau RCV1. Namun, analisis kritis terhadap pendekatan ini mengungkapkan adanya hambatan komputasi yang tinggi dan kebutuhan akan ribuan data latih untuk mencapai konvergensi model yang optimal. Dalam konteks Inrenbang, di mana *dataset* yang tersedia terbatas pada sampel internal dari atasan magang, penerapan model yang terlalu kompleks berisiko mengalami *overfitting* dan sulit untuk diimplementasikan pada infrastruktur server yang sederhana. Oleh karena itu, pendekatan *problem transformation* dengan algoritma pembelajaran mesin konvensional dinilai lebih pragmatis dan efisien (Duan et al., 2022).

Penelitian selanjutnya yang relevan membahas perbandingan metode *problem transformation* pada domain spesifik seperti medis dan klasifikasi berita. Hasil studi menunjukkan bahwa *Classifier Chains* sering kali mengungguli *Binary Relevance* ketika terdapat korelasi positif yang kuat antar label. Namun, keterbatasan dari studi-studi ini adalah kurangnya pembahasan mengenai dampak urutan rantai (*chain order*) pada dataset yang tidak seimbang (*imbalanced dataset*), yang merupakan karakteristik umum dari data administrasi kepegawaian. Selain itu, banyak penelitian sebelumnya tidak mengintegrasikan model klasifikasi tersebut secara langsung ke dalam antarmuka *chatbot real-time*, melainkan hanya sebatas evaluasi performa pada data statis. Hal ini meninggalkan celah mengenai bagaimana latensi model mempengaruhi pengalaman pengguna secara langsung saat berinteraksi dengan sistem *helpdesk* (Arslan & Cruz, 2024).

Berdasarkan tinjauan tersebut, teridentifikasi adanya *research gap* yang signifikan, yaitu kurangnya studi yang secara spesifik menerapkan dan membandingkan metode *Classifier Chains* dan *Binary Relevance* untuk klasifikasi *multi-label* pada domain administrasi kepegawaian dengan batasan *dataset* lokal yang tertutup. Penelitian ini menawarkan perspektif baru dengan merancang model yang tidak hanya akurat secara statistik tetapi juga dioptimalkan untuk *knowledge base* yang terbatas, memastikan jawaban *chatbot* tetap terkontrol dan sesuai dengan regulasi internal tanpa terdistraksi oleh informasi eksternal. Pendekatan ini berbeda karena berfokus pada maksimalisasi utilitas *small data* melalui rekayasa metode klasifikasi, memberikan kontribusi mendalam mengenai bagaimana institusi dengan sumber daya data terbatas tetap dapat membangun sistem AI yang cerdas dan responsif (Du, 2024).

Kontribusi penelitian ini terhadap ilmu pengetahuan terletak pada pengayaan literatur mengenai penerapan *Natural Language Processing* (NLP) untuk kasus *multi-label* di lingkungan *low-resource setting*. Hasil komparasi antara *Classifier Chains* dan *Binary Relevance* akan memberikan bukti empiris mengenai metode mana yang lebih resisten terhadap keterbatasan data dalam bahasa Indonesia, khususnya ragam bahasa administratif. Secara teoritis, penelitian ini memperkuat pemahaman mengenai pentingnya menangani korelasi label dalam teks pendek untuk meningkatkan presisi sistem tanya jawab. Temuan ini diharapkan dapat menjadi referensi bagi peneliti selanjutnya yang menghadapi kendala serupa dalam pengolahan data teks yang bersifat spesifik dan rahasia (Ariyana, 2025).

Sebagai penutup, kontribusi praktis dari penelitian ini sangat signifikan bagi sektor administrasi publik dan manajemen kepegawaian. Bagi Inrenbang, sistem yang dihasilkan akan mengakselerasi respon *helpdesk*, mengurangi waktu tunggu pegawai, dan meminimalisir *human error* dalam pemberian informasi, yang secara langsung mendukung efisiensi birokrasi. Bagi pemangku kebijakan, penelitian ini memberikan rekomendasi teknis mengenai kelayakan adopsi teknologi AI sederhana namun efektif untuk digitalisasi layanan internal, tanpa perlu investasi infrastruktur yang masif. Pada akhirnya, penelitian ini berdampak pada terciptanya ekosistem kerja yang lebih responsif dan modern, membuktikan bahwa inovasi teknologi dapat diterapkan secara efektif dari skala unit kerja terkecil sekalipun (Saprudin, 2024).

1.2 Identifikasi Masalah

Pelayanan *helpdesk* di lingkungan Inrenbang saat ini menghadapi tantangan operasional yang serius akibat ketergantungan pada penanganan manual dalam merespons pertanyaan pegawai. Volume interaksi yang tinggi sering kali menyebabkan terjadinya antrean respons yang panjang, sehingga menghambat aliran informasi penting terkait administrasi kepegawaian. Situasi ini diperburuk oleh ketidakkonsistenan jawaban yang diberikan oleh staf

yang berbeda, yang memicu kebingungan di kalangan pengguna layanan. Fenomena keterlambatan ini menunjukkan adanya urgensi untuk beralih dari sistem konvensional menuju sistem otomasi cerdas yang mampu bekerja secara *real-time* untuk menjaga efisiensi kinerja organisasi (Fridayanto & Wibowo, 2025).

Permasalahan teknis utama yang menjadi penghambat pengembangan otomasi di Inrenbang adalah karakteristik pertanyaan pegawai yang cenderung kompleks dan mengandung banyak intensi (*multi-label*). Seorang pegawai sering kali mengajukan satu kalimat tanya yang mencakup dua topik sekaligus, seperti menanyakan sisa cuti tahunan dan prosedur klaim tunjangan kesehatan secara bersamaan. Sistem klasifikasi standar yang hanya mampu mendeteksi label tunggal akan gagal memahami konteks ini, sehingga jawaban yang dihasilkan menjadi parsial atau tidak relevan. Kegagalan dalam memetakan seluruh intensi pengguna ini menjadi titik lemah utama yang harus segera diselesaikan (Cahyadi et al., 2025)

Kendala signifikan lainnya terletak pada keterbatasan sumber data (*resource constraint*) yang tersedia untuk melatih model kecerdasan buatan. Pengembangan *chatbot* ini tidak dapat memanfaatkan *dataset* global atau korpus internet karena harus mematuhi regulasi internal dan spesifikasi jawaban yang ditetapkan oleh institusi. Data latih hanya bersumber dari sampel terbatas yang disediakan oleh atasan magang, sehingga model harus mampu belajar secara optimal dari data yang sedikit (*small data*). Hal ini menciptakan kesulitan tersendiri dalam mencapai akurasi tinggi tanpa risiko *overfitting*, berbeda dengan model yang dilatih pada *big data* (Kumar et al., 2024).

Dampak negatif yang akan timbul jika permasalahan klasifikasi *multi-label* dan keterbatasan data ini tidak ditangani adalah stagnasi kualitas pelayanan publik di Inrenbang. *Chatbot* yang tidak akurat akan menurunkan kepercayaan pengguna, menyebabkan pegawai kembali membebani staf manusia dengan pertanyaan berulang, yang pada akhirnya menggagalkan tujuan efisiensi digitalisasi. Selain itu, ketidakmampuan sistem dalam memberikan respons yang presisi berpotensi menimbulkan kesalahan interpretasi kebijakan kepegawaian, yang dapat berujung pada kerugian administratif bagi pegawai maupun institusi. Oleh karena itu, solusi algoritmik yang tepat sangat dibutuhkan untuk mencegah degradasi layanan ini (Kocharyan, 2022).

Tinjauan terhadap literatur terdahulu menunjukkan bahwa mayoritas penelitian *chatbot* layanan publik masih terpaku pada metode klasifikasi *single-label* seperti *Support Vector Machine* atau Naïve Bayes. Meskipun metode tersebut efektif untuk pertanyaan sederhana, kinerjanya merosot tajam ketika dihadapkan pada kalimat majemuk yang umum dalam komunikasi birokrasi. Kesenjangan ini menunjukkan bahwa pendekatan konvensional tidak

lagi memadai untuk menangani kompleksitas bahasa alami dalam konteks administrasi modern. Belum banyak studi yang mengeksplorasi solusi khusus untuk skenario *multi-intent* pada domain tertutup dengan data yang sangat terbatas (Caldarini et al., 2022).

Lebih lanjut, terdapat kebutuhan mendesak untuk membandingkan efektivitas metode *problem transformation* spesifik, yaitu *Classifier Chains* dan *Binary Relevance*, dalam konteks data administrasi berbahasa Indonesia. Penelitian sebelumnya sering kali mengabaikan analisis perbandingan kedua metode ini pada *dataset* kecil yang tidak seimbang (*imbalanced dataset*), padahal karakteristik tersebut sangat dominan dalam data kepegawaian. *Binary Relevance* unggul dalam kecepatan namun mengabaikan korelasi label, sedangkan *Classifier Chains* memperhitungkan ketergantungan antar label namun lebih kompleks. Belum adanya kepastian mengenai metode mana yang paling optimal untuk kasus Inrenbang menjadi celah penelitian yang krusial (Ding et al., 2022).

Berdasarkan uraian di atas, inti permasalahan penelitian ini mengerucut pada bagaimana merancang model klasifikasi teks yang mampu menangani input *multi-label* secara akurat di tengah keterbatasan *dataset* internal Inrenbang. Masalah utamanya adalah ketidaktahuan mengenai performa komparatif antara metode *Classifier Chains* dan *Binary Relevance* dalam meningkatkan relevansi jawaban *chatbot* administrasi. Penelitian ini menjadi penting untuk menjawab tantangan teknis tersebut guna menciptakan asisten virtual yang cerdas, responsif, dan sesuai dengan kebutuhan operasional spesifik tanpa bergantung pada sumber daya data eksternal yang masif (Nofriani & Kurniawan, 2021).

1.3 Rumusan Masalah

Berdasarkan identifikasi masalah yang diperoleh, maka rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana merancang model klasifikasi teks multi-label menggunakan metode *Classifier Chains* dengan *Logistic Regression* dan *TF-IDF Vectorizer* yang dapat mengklasifikasikan pertanyaan pegawai ke dalam lebih dari satu kategori layanan kepegawaian secara akurat?
2. Bagaimana mengimplementasikan model klasifikasi *multi-label* tersebut ke dalam sistem chatbot yang terintegrasi untuk melayani pertanyaan administrasi kepegawaian di lingkungan INRENBANG?
3. Bagaimana mengukur efektivitas sistem chatbot berbasis klasifikasi multi-label dalam mengakselerasi waktu respon helpdesk dibandingkan dengan penanganan manual?

1.4 Tujuan Penelitian

Penelitian ini bertujuan mengembangkan sistem Chatbot untuk membantu Pertanyaan dari semua Para Admin Satuan Kerja :

1. Merancang model klasifikasi teks multi-label menggunakan metode *Classifier Chains* dengan algoritma *Logistic Regression* dan *TF-IDF Vectorizer* yang mampu mengklasifikasikan pertanyaan pegawai ke dalam lima kategori layanan kepegawaian (Absensi, Tunjangan Kinerja, Cuti, Cacah Jiwa, dan Tugas Belajar) dengan tingkat akurasi yang optimal.
2. Mengimplementasikan model klasifikasi multi-label ke dalam sistem chatbot terintegrasi berbasis web yang dapat digunakan untuk melayani pertanyaan administrasi kepegawaian di lingkungan INRENBANG secara otomatis dan real-time
3. Mengukur dan mengevaluasi efektivitas sistem chatbot berbasis klasifikasi multi-label dalam mengakselerasi waktu respon helpdesk dengan membandingkan kecepatan respons sistem terhadap penanganan manual konvensional

1.5 Batasan Masalah

Pembatasan masalah Pembatasan masalah diperlukan untuk memperjelas ruang lingkup topik yang diteliti sehingga penelitian dapat lebih terarah dan fokus pada isu utama yang dibahas:

1. Penelitian ini dibatasi pada perancangan model klasifikasi teks *multi-label* untuk domain administrasi kepegawaian INRENBANG dengan lima kategori label, yaitu Absensi, Tunjangan Kinerja, Cuti, Cacah Jiwa, dan Tugas Belajar. Metode yang digunakan terbatas pada Classifier Chains dengan *Logistic Regression* dan *TF-IDF Vectorizer*, tanpa melibatkan pendekatan deep learning
2. Dataset yang digunakan berupa pertanyaan pegawai dari helpdesk INRENBANG dengan jumlah ±150 sampel. Evaluasi difokuskan pada performa model menggunakan metrik klasifikasi multi-label dan pengukuran waktu respon sistem, tanpa melakukan user acceptance testing dalam skala besar.
3. Variabel independen dalam penelitian ini adalah fitur teks hasil ekstraksi TF-IDF, sedangkan variabel dependen adalah hasil klasifikasi multi-label berupa kombinasi lima kategori layanan. Variabel evaluasi meliputi F1-Score, Hamming Loss, Subset Accuracy, dan waktu respon sistem dalam milidetik.

1.6 Kontribusi Penelitian

Penelitian ini memberikan kontribusi dalam pengembangan dan penerapan praktis metode *Classifier Chains* dengan *Logistic Regression* dan *TF-IDF Vectorizer* dalam perancangan chatbot helpdesk administrasi kepegawaian:

1. Mengembangkan sistem chatbot otomatis untuk klasifikasi pertanyaan pegawai ke dalam lima kategori layanan kepegawaian (Absensi, Tunjangan Kinerja, Cuti, Cacah Jiwa, dan Tugas Belajar), mempercepat waktu respon helpdesk dibandingkan penanganan manual.
2. Menerapkan metode Classifier Chains dengan Logistic Regression untuk klasifikasi teks multi-label, memungkinkan sistem mengenali lebih dari satu konteks dalam satu pertanyaan sehingga menghasilkan respons yang lebih komprehensif dan akurat.
3. Menggunakan TF-IDF Vectorizer untuk ekstraksi fitur teks berbahasa Indonesia dengan pendekatan n-gram, meningkatkan kemampuan model dalam menangkap pola kata dan frasa yang relevan dengan domain administrasi kepegawaian. Selain itu, penelitian ini membandingkan performa Binary Relevance dan Classifier Chains dalam hal F1-Score dan Hamming Loss, memberikan wawasan tentang metode terbaik untuk klasifikasi multi-label pada domain spesifik.
4. Mengimplementasikan sistem chatbot berbasis web yang terintegrasi dengan model machine learning, mendukung penerapan Sistem Pemerintahan Berbasis Elektronik

(SPBE) dan dapat menjadi prototipe bagi instansi pemerintah lainnya dalam mengembangkan layanan helpdesk otomatis yang responsif dan efisien.


UNIVERSITAS NASIONAL