

ABSTRAK

Permasalahan utama yang dihadapi oleh PT PLN (Persero) ULP Padalarang UP3 Cimahi UID Jawa Barat adalah tingginya tingkat kesalahan dalam klasifikasi keluhan pelanggan. Kondisi ini berdampak pada keterlambatan penanganan, alokasi unit teknis yang tidak tepat, serta menurunnya kepuasan pelanggan. Solusi yang telah tersedia seperti Contact Center 123 dan aplikasi PLN Mobile masih memiliki keterbatasan dalam hal aksesibilitas, kecepatan respons, dan kemampuan memahami konteks bahasa pelanggan.

Untuk menjawab tantangan tersebut, penelitian ini mengembangkan sistem chatbot berbasis *Natural Language Processing* (NLP) yang mampu memahami dan mengklasifikasikan intent keluhan pelanggan secara otomatis dan akurat. Pemanfaatan chatbot dipilih karena mampu memberikan layanan interaktif secara *real-time* tanpa intervensi manusia, sehingga mempercepat proses klasifikasi dan penanganan keluhan. Sementara itu, pendekatan NLP digunakan karena efisien dalam memahami teks berbahasa alami dan dapat dioperasikan secara lokal, menjadikannya lebih terjangkau dan terintegrasi dibandingkan model *Large Language Model* (LLM) yang memerlukan komputasi tinggi dan infrastruktur khusus.

Lima model NLP dibandingkan dalam penelitian ini, yaitu IndoBERT, mBERT, TF-IDF + Naive Bayes, Word2Vec + SVM, dan FastText. Dataset dikumpulkan dari layanan PLN *Mobile* dan *Contact Center* 123, mencakup lima kategori keluhan utama: Pasang Baru, Perubahan Daya, Gagal Input Token, kWh Periksa, dan Listrik Padam. Evaluasi performa model dilakukan menggunakan 5-fold cross-validation dan sejumlah metrik: akurasi, presisi, *recall*, *F1-Score*, MCC, dan Kappa Cohen. Selanjutnya, untuk menganalisis perbedaan performa antar model secara statistik, digunakan uji Friedman dan post-hoc Nemenyi.

Hasil penelitian menunjukkan bahwa model IndoBERT memberikan performa terbaik, dengan akurasi 95,83%, *F1-Score* makro 96,02%, dan MCC 93,45%. Uji Friedman menghasilkan nilai $p = 0,0012$, yang mengindikasikan adanya perbedaan signifikan antar model. Chatbot yang dikembangkan berhasil diintegrasikan dengan aplikasi PLN *Mobile* menggunakan *framework Flask*, dan mampu mengklasifikasikan intent keluhan secara *real-time* dengan *confidence threshold* di atas 0,8. Penelitian ini juga menghasilkan benchmark dataset klasifikasi keluhan utilitas dalam Bahasa Indonesia serta kerangka evaluasi model klasifikasi yang dapat digunakan dalam penelitian lanjutan.

Kata kunci— chatbot, PLN *Mobile*, *artificial intelligence*, BERT, PT PLN, model transformer, kepuasan pelanggan, *deep learning*, interaksi pelanggan, ULP Padalarang, UP3 Cimahi, UID Jawa Barat

ABSTRACT

The main issue faced by PT PLN (Persero) ULP Padalarang UP3 Cimahi UID West Java is the high error rate in classifying customer complaints. This condition results in delayed handling, inaccurate allocation to technical units, and a decline in customer satisfaction. Existing solutions such as Contact Center 123 and the PLN Mobile application still have limitations in terms of accessibility, response speed, and the ability to understand the context of customer language.

To address this challenge, this study develops a chatbot system based on Natural Language Processing (NLP) capable of automatically and accurately understanding and classifying customer complaint intents. The chatbot approach is selected for its ability to provide real-time interactive service without human intervention, thereby accelerating the complaint classification and handling process. NLP is chosen over Large Language Models (LLM) due to its efficiency in processing natural language text and its suitability for local deployment, making it more cost-effective and integrable within PLN's infrastructure.

Five NLP models are compared in this study: IndoBERT, mBERT, TF-IDF + Naive Bayes, Word2Vec + SVM, and FastText. The dataset was collected from PLN Mobile and Contact Center 123 services, comprising five main complaint categories: New Installation, Power Change, Token Input Failure, kWh Check, and Power Outage. Model performance was evaluated using 5-fold cross-validation and multiple metrics including accuracy, precision, recall, F1-Score, Matthews Correlation Coefficient (MCC), and Cohen's Kappa. To determine whether performance differences among models were statistically significant, Friedman test and Nemenyi post-hoc analysis were applied.

The results indicate that the IndoBERT model outperformed the others, achieving 95.83% accuracy, 96.02% macro F1-Score, and 93.45% MCC. The Friedman test yielded a p-value of 0.0012, indicating statistically significant differences across models. The chatbot developed using the Flask framework was successfully integrated into the PLN Mobile application and is capable of real-time intent classification with a confidence threshold above 0.8. This study also contributes a benchmark dataset for Indonesian utility complaint classification and an evaluation framework applicable for future research.

Keywords— chatbot, PLN Mobile, artificial intelligence, BERT, PT PLN, transformer model, customer satisfaction, deep learning, customer interaction, ULP Padalarang, UP3 Cimahi, UID Jawa Barat