

Deteksi Kelengkapan Laporan Aduan pada Data Aduan Masyarakat di Pemerintah Provinsi Daerah Khusus Ibukota Jakarta = Detection of Complaint Reports Completeness on Public Complaint Data in the Provincial Government of the Special Capital Region of Jakarta

Hamam Wulan Ayu, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920543029&lokasi=lokal>

Abstrak

Pemerintah Provinsi DKI Jakarta telah mengembangkan sistem Cepat Respon Masyarakat untuk mendukung partisipasi masyarakat. Masyarakat DKI Jakarta dapat menyampaikan keluhan non-darurat di 14 kanal dan memantau proses penyelesaiannya di portal yang tersedia. Namun, tidak semua aduan yang disampaikan masyarakat mempunyai keterangan yang lengkap. Beberapa aduan tidak menyertakan deskripsi, lokasi, maupun waktu kejadian. Hal ini menyulitkan petugas Operasi Perangkat Daerah (OPD) dalam melakukan penyelesaian aduan. Dalam menangani permasalahan tersebut, diperlukan suatu sistem yang dapat digunakan untuk mendeteksi kelengkapan suatu aduan dari masyarakat, sehingga dapat secara otomatis mengingatkan masyarakat untuk membuat laporan yang lengkap. Penelitian ini membahas performa model yang dapat mendeteksi kelengkapan keterangan laporan aduan. Rangkaian metode yang diterapkan adalah deteksi kelengkapan keterangan deskripsi, deteksi kelengkapan keterangan lokasi, dan deteksi kelengkapan keterangan waktu. Deteksi kelengkapan keterangan deskripsi diuji dengan pendekatan machine learning menggunakan fitur Term Frequency-Inverse Document Frequency (TF-IDF) dan N-gram. Algoritma machine learning yang diimplementasikan adalah Multinomial Na ve Bayes (MNB), Support Vector Machine (SVM), dan Random Forest Decision Tree (RFDT). Deteksi kelengkapan keterangan lokasi dan waktu diuji menggunakan sistem Named Entity Recognition (NER) Wirawan (2020) dengan model pre-trained Bidirectional Encoder Representations from Transformers (BERT) dan basis data NERGRIT. Ekstraksi kata kunci keterangan deskripsi dilakukan menggunakan pendekatan LIME explainer. Hasil eksperimen menunjukkan bahwa algoritma RFDT menggunakan fitur TF dan word unigram dengan balancing data menggunakan Synthetic Minority Over-sampling Technique (SMOTE) menghasilkan performa tertinggi untuk klasifikasi kelengkapan keterangan deskripsi laporan aduan dengan nilai makro F1 sebesar 0.7154. Deteksi kelengkapan keterangan lokasi dan waktu menggunakan model NER pre-trained BERT menghasilkan nilai makro F1 sebesar 0.8634 dan 0.6252. Hasil pemetaan kata kunci keterangan deskripsi, lokasi dan waktu dievaluasi dengan metrik cosine similarity secara berurutan menghasilkan nilai sebesar 0.5890, 0.6252 dan 0.8487.

.....The Provincial Government of DKI Jakarta has developed a Cepat Respon Masyarakat system to support community participation and optimization of government performance. DKI Jakarta citizens can submit non-emergency complaints in 14 channels integrated with the Cepat Respon Masyarakat and monitor the resolution process on the available portal. However, not all complaints submitted by the public have complete information. Some complaints do not include a description, location, or time of occurrence. This will make it difficult for the Organisasi Perangkat Daerah (OPD) officers to resolve them. To deal with this problem, we need a system that can detect the completeness of a complaint from the public to automatically remind them to make a complete report. This study discusses the performance of models that can detect complaint report completeness. The methods applied are detection of completeness of description

information, detection of completeness of location information, and detection of completeness of time information. Detection of completeness of descriptions was tested using the machine learning approach with the Term Frequency-Inverse Document Frequency (TF-IDF) and N-gram as the features. The implemented machine learning algorithms are Multinomial Na ve Bayes (MNB), Support Vector Machine (SVM), and Random Forest Decision Tree (RDFT). Detection of completeness of location and time descriptions was tested using Named Entity Recognition (NER) system by Wirawan (2020) with pre-trained Bidirectional Encoder Representations from Transformers (BERT) model dan NERGRIT as data basis. Description's keyword extraction was performed by LIME explainer implementation. The experimental results show that the RFDT algorithm with TF and word unigram features and data balancing using Synthetic Minority Over-sampling Technique (SMOTE) produced the highest performance for the classification of the completeness of the description with F1 macro value of 0.7154. Classification of completeness of location and time information using the pre-trained BERT model resulted in F1 macro values of 0.8634 and 0.6252. The results of the detection of description keywords, location keywords, and time keywords were evaluated using the cosine similarity metric, respectively, yielding a value of 0.863448, 0.6252, and 0.8487.