

INTISARI

ANALISIS SENTIMEN BERBASIS ASPEK MENGGUNAKAN SUPERVISED LEARNING PADA SOSIAL MEDIA E-GOVERNMENT (STUDI KASUS DIREKTORAT JENDERAL IMIGRASI KEMENTERIAN HUKUM DAN HAM REPUBLIK INDONESIA)

Oleh

Fathiyarizq Mahendra Putra

22/500625/PPA/06359

Pemerintah dalam sektor Pusat, K/L maupun Pemerintahan Daerah, membuat akun sosial media untuk dapat mendekatkan diri kepada warganet, serta merespon masukan dan keluhan masyarakat, namun cukup sulitnya memahami substansi dan aspek-aspek dari opini tentang layanan yang diberikan oleh warganet, di dalam opini tersebut sesuai aspek performa dan kinerja serta layanan atau hanya sebuah opini lainnya, penelitian ini mengusulkan model algoritma untuk analisis sentimen dan aspek opini dalam menilai kebijakan dan layanan Direktorat Jenderal Imigrasi Kementerian Hukum dan HAM RI melalui opini pengguna di Twitter. Fokus penelitian adalah mengembangkan model terbaik dari beberapa algoritma *supervised learning* untuk klasifikasi aspek dan sentimen menggunakan dataset tweet dari pengguna layanan dan opini publik.

Rancangan terdiri dari pengumpulan data dan filtering data, preprocessing data, *Word Embedding*, dan *splitting data* menggunakan *K-Fold Cross Validation*. Pada pengumpulan dan labeling terdapat ketidakseimbangan data antar label sehingga diperlukan SMOTE untuk menangani permasalahan tersebut, tahapan klasifikasi dengan algoritma *supervised learning*, dan Evaluasi pengujian. Hasil penelitian menunjukkan bahwa teknik Word2Vec, khususnya Skip-gram, memberikan performa lebih baik dibandingkan CBOW. Pada aspek "pelayanan", Random Forest (Skip-gram) mencapai akurasi 98.36%, presisi 99.79%, recall 97.78%, dan F1-Score 98.78%. Untuk aspek "pelayanan", K-Nearest Neighbor memberikan akurasi 62.167%, presisi 62.051%, recall 98.641%, dan F1-Score 76.18%. Pada aspek "kebijakan", SVM menghasilkan akurasi 77.417%, namun dengan keterbatasan interpretasi akibat ketidakseimbangan data.

Penelitian ini menekankan pentingnya teknik preprocessing seperti stemming, penyeimbangan data melalui SMOTE, dan ekstraksi fitur menggunakan Skip-gram dalam meningkatkan performa model klasifikasi pada dataset tidak terstruktur dari media sosial, khususnya untuk analisis opini layanan publik.

Kata Kunci: Analisis Sentimen, *aspect-based sentiment*, *machine learning*, *supervised learning*

ABSTRACT

ASPECT-BASED SENTIMENT ANALYSIS USING SUPERVISED LEARNING IN E-GOVERNMENT SOCIAL MEDIA (CASE STUDY OF THE DIRECTORATE GENERAL OF IMMIGRATION MINISTRY OF LAW AND HUMAN RIGHTS OF THE REPUBLIC OF INDONESIA)

By

Fathiyarizq Mahendra Putra

22/500625/PPA/06359

The government in the central sector, ministries/institutions and local governments, create social media accounts to be able to get closer to citizens, as well as respond to them. social media accounts to be able to get closer to citizens, as well as respond to public input and complaints. input and complaints from the public, but it is quite difficult to understand the substance and aspects of opinions about services provided by netizens, in accordance with the performance and aspects of performance and service or just another opinion. just another opinion. This research proposes an algorithm model for sentiment analysis and opinion aspects in assessing the policies and services of the Directorate General of Immigration of the Ministry of Law and Human Rights of the Republic of Indonesia through user opinions on Twitter. The focus of the research is to develop the best model from several supervised learning algorithms for aspect and sentiment classification using a dataset of tweets from service users and public opinion.

The design of this research consists of data collection and data filtering, data preprocessing, Word Embedding, and data splitting using K-Fold Cross Validation. In collecting and labeling there is an imbalance of data between labels so SMOTE is needed to handle this problem, classification stages with supervised learning algorithms, and evaluation testing. Results demonstrate that the Word2Vec technique, particularly Skip-gram, outperformed CBOW. For the "service" aspect, Random Forest (Skip-gram) achieved 98.36% accuracy, 99.79% precision, 97.78% recall, and 98.78% F1-Score. K-Nearest Neighbor yielded 62.167% accuracy, 62.051% precision, 98.641% recall, and 76.18% F1-Score. The "policy" aspect, analyzed using SVM, generated 77.417% accuracy, with interpretation limitations due to data imbalance.

The study underscores the critical importance of preprocessing techniques like stemming, data balancing through SMOTE, and feature extraction using Skip-gram in enhancing classification model performance on unstructured social media datasets, particularly for public service opinion analysis.

Keywords: Sentiment Analysis, aspect-based sentiment, machine learning, supervised learning.