ABSTRAK

PERBANDINGAN AKURASI METODE *LEXICON BASED* DAN METODE *FEATURE EXTRACTION* MENGGUNAKAN ALGORITMA *SUPPORT VECTOR MACHINE* (SVM) PADA ANALISIS SENTIMEN LAYANAN AKADEMIK UIN SUNAN GUNUNG DJATI BANDUNG

Oleh :

Ifany Dewi Tustianti

1197050049

Penelitian ini menggunakan metode Lexicon based dan Feature extraction dengan algoritma Support Vector Machine (SVM) dalam bahasa pemrograman Python untuk menganalisis sentimen terhadap layanan akademik UIN Sunan Gunung Djati Bandung. Dalam lima percobaan, metode Lexicon based mencapai akurasi 95%, sementara metode Feature extraction, dengan Term Presence, BoW, dan TF-IDF, memiliki akurasi masing-masing 84%, 84%, dan 82%. Hasil klasifikasi menunjukkan kecenderungan masyarakat memiliki pandangan netral terhadap layanan tersebut. Validasi silang 10-fold menunjukkan metode Lexicon based memiliki akurasi rata-rata 96,20%, sementara metode Feature extraction memiliki akurasi rata-rata masing-masing 92,30%, 92,20%, dan 89,90%. Penelitian ini menyimpulkan bahwa metode Lexicon based lebih konsisten dan efektif dalam mengklasifikasikan sentimen, dengan BoW dianggap sebagai teknik ekstraksi fitur paling efektif dalam Feature extraction. Hasilnya menegaskan bahwa metode Lexicon based dapat diandalkan untuk analisis sentimen layanan akademik UIN Sunan Gunung Djati Bandung.

Kata Kunci: Analisis Sentimen, *Lexicon based* , *Feature extraction*, *Support Vector Machine* (SVM), Akurasi, *CRISP-DM*

ABSTRACT

COMPARISON OF ACCURACY OF LEXICON BASED METHOD AND FEATURE EXTRACTION METHOD USING SUPPORT VECTOR MACHINE (SVM) ALGORITHM ON SENTIMENT ANALYSIS OF ACADEMIC SERVICES UIN SUNAN GUNUNG DJATI BANDUNG

By :

Ifany Dewi Tustianti

1197050049

This research uses Lexicon based and Feature extraction methods with the Support Vector Machine (SVM) algorithm in the Python programming language to analyze sentiment towards the academic services of UIN Sunan Gunung Djati Bandung. In five experiments, the Lexicon based method achieved 95% accuracy, while the Feature extraction method, with Term Presence, BoW, and TF-IDF, had 84%, 84%, and 82% accuracy respectively. The classification results show a tendency for people to have a neutral view of the service. The 10-fold cross-validation shows that the Lexicon based method has an average accuracy of 96.20%, while the Feature extraction method has an average accuracy of 92.30%, 92.20%, and 89.90%, respectively. This study concludes that the Lexicon based method is more consistent and effective in classifying sentiment, with BoW being considered the most effective feature extraction technique in Feature extraction. The results confirm that the Lexicon based method is reliable for sentiment analysis of academic services of UIN Sunan Gunung Djati Bandung.

Keywords: Sentiment Analysis, Lexicon based , Feature extraction, Support Vector Machine (SVM), Accuracy, CRISP-DM