

## ABSTRAK

Nama	:	Muhamad Rizky Ardiansyah
NIM	:	41520110050
Program Studi	:	Teknik Informatika
Judul Laporan Skripsi	:	Analisis Sentimen Masyarakat Terhadap Film “Dirty Vote” di YouTube Menggunakan Model SVM (Support Vector Machine) dan LSTM (Long Short-Term Memory)
Pembimbing	:	Suhendra S. Kom., M. Kom.

Film dokumenter "Dirty Vote" yang tersedia di kanal YouTube PSHK Indonesia, telah menjadi salah satu video yang memperoleh perhatian dan reaksi yang beragam di *platform* YouTube. Analisis sentimen dilakukan untuk mengetahui sentimen masyarakat dalam memberikan komentar terhadap film dokumenter ini. Dalam menganalisis sentimen masyarakat, digunakan *InSet Lexicon* sebagai kamus untuk proses pelabelan otomatis bahasa Indonesia, *SVM* dan *LSTM* sebagai algoritma klasifikasi sentimen dan *SMOTE* untuk mengatasi *overfitting*. Penelitian ini melakukan perbandingan hasil pelabelan dan model di tiap tahap pra-pemrosesan teks. Hasil penelitian menunjukkan penggunaan *InSet Lexicon* untuk pelabelan pada tahap terakhir di pra-pemrosesan *stemming* menghasilkan distribusi sentimen yang paling timpang, sentimen positif sebesar 20,26% dan negatif mencapai 79,74%. Selain itu, hasil evaluasi dari kinerja *SVM* mencatat akurasi tertinggi sebesar 97,58% menggunakan data di proses *stemming* dengan *test size* 10% dan *random state* 1, sementara *LSTM* mencatat akurasi tertinggi sebesar 96,3% dan nilai *loss function* yaitu 0,098 dengan *batch size* 16 dengan menggunakan data ditahap *stemming*. Penggunaan metode *SMOTE* untuk mengatasi *overfitting* menunjukkan hasil yang baik untuk model *SVM* sedangkan hasil yang negatif untuk *LSTM*.

Kata kunci: Analisis sentimen, *SVM*, *LSTM*, *SMOTE*, *Lexicon*

## ABSTRACT

Nama : Muhamad Rizky Ardiansyah  
NIM : 41520110050  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : *Analysis of Public Sentiment towards the Movie "Dirty Vote" on YouTube Using SVM Model (Support Vector Machine) and LSTM (Long Short-Term Memory) models.*  
Pembimbing : Suhendra S. Kom., M. Kom.

*The documentary "Dirty Vote" on PSHK Indonesia's YouTube channel has become one of the videos that has gained attention and mixed reactions on the YouTube platform. Sentiment analysis was conducted to find out people's sentiments in commenting on this documentary. In analyzing public sentiment, InSet Lexicon is used as a dictionary for the automatic labeling process of Indonesian language, SVM and LSTM as sentiment classification algorithms and SMOTE to overcome overfitting. This research compares the labeling results and models at each stage of text pre-processing. The results showed that the use of InSet Lexicon for labeling at the last stage of stemming pre-processing resulted in the most unequal sentiment distribution, with positive sentiment accounting for 20,26% and negative sentiment accounting for 79,74%. In addition, the evaluation results of SVM performance recorded the highest accuracy of 97,58% using data in the stemming process with test size 10% and random state 1, while LSTM recorded the highest accuracy of 96,3% and loss function value of 0,098 with batch size 16 using data in the stemming stage. The use of the SMOTE method to overcome overfitting showed good results for the SVM model while negative results for the LSTM.*

*Keywords: Sentiment Analysis, SVM, LSTM, SMOTE, Lexicon*