

## ABSTRAKSI

### **STUDI PERBANDINGAN METODE EKSTRAKSI FITUR, SELEKSI FITUR, DAN KLASIFIKASI PADA ANALISIS SENTIMEN SELAMA PEMILIHAN UMUM INDONESIA 2019**

Oleh

GAMA CANDRA TRI KARTIKA

15/378060/PA/16535

Data dari media sosial seperti Twitter dan situs portal berita seperti Line-Today merupakan data yang objektif untuk diolah dikarenakan penggunaanya yang sangat aktif dalam menyampaikan pendapatnya di kedua media sosial dan portal berita tersebut. Namun perbandingan metode dalam analisis sentimen belum banyak diterapkan dalam domain Bahasa Indonesia dan dataset Twitter dan Line-Today.

Tujuan dari penelitian ini adalah untuk mengetahui nilai terbaik dari metode ekstraksi fitur, seleksi fitur, dan metode pengklasifikasian serta pemilihan kombinasi yang terbaik dari proses seleksi fitur dan ekstraksi fitur, terhadap nilai analisis sentimen dengan berbagai metode pengklasifikasian.

Hasil penelitian ini adalah pada dataset Line-Today, pengklasifikasi terbaik pada masing-masing rata-rata semua nilai adalah Extra Tree (akurasi = 84.15 %), dan Extra Tree Classifier (*f1 score* = 78.44 %). Kemudian pada dataset Twitter, pengklasifikasi terbaik pada masing-masing rata-rata semua nilai adalah Extra Tree (akurasi = 81.25 %) dan Extra Tree Classifier (*f1 score* = 78.45 %).

Kata-kata kunci : analisis sentimen, seleksi fitur, ekstraksi fitur, klasifikasi teks

## ABSTRACT

### **A COMPARATIVE STUDY OF FEATURES EXTRACTION, FEATURES SELECTION, AND CLASSIFIER METHODS ON SENTIMENT ANALYSIS DURING INDONESIAN ELECTION 2019**

By

GAMA CANDRA TRI KARTIKA  
15/378060/PA/16535

Data from social media such as Twitter and online news portal such as Line-Today are objective data to be processed because users are very active in expressing their opinions on both social media and online news portal. But the comparison of methods in sentiment analysis has not been widely applied in the Indonesian language, Twitter and Line-Today dataset .

The purpose of this study was to determine the best value of feature extraction methods, feature selection methods, and classification methods and the selection of the best combination of feature selection and feature extraction processes, to the value of sentiment analysis with various classification methods.

The results of this study on the Line-Today dataset, the best classifiers for each average of all values are Extra Trees (accuracy = 84.15 %), and Extra Tree Classifiers (f1 score = 78.44 %). Then on the Twitter dataset, the best classifiers in each average of all values are Extra Trees (accuracy = 81.25 %), and Extra Tree Classifier ( f1 score = 78.45 %).

Keywords: sentiment analysis, feature selection, feature extraction, text classification