

## ABSTRAK

Pemilu Presiden dan Wakil Presiden Indonesia merupakan salah satu ajang perwujudan demokrasi di Indonesia yang akan memilih Presiden dan Wakil Presiden Indonesia setiap 5 tahun sekali. Berdasarkan data, jumlah Daftar Pemilih Tetap (DPT) yang melebihi 200 juta orang, menjadikan pemanfaatan internet penting bagi pasangan kandidat untuk mendapatkan dukungan. Salah satu situs web yang memiliki banyak pengunjung di Indonesia adalah YouTube. Tayangan debat kepresidenan di YouTube memantik ruang diskusi publik di dalam kolom komentar video, di mana pada kolom komentar tersebut terdapat opini positif dan negatif dari masyarakat.

Pada penelitian ini, dilakukan analisis sentimen pada setiap pasangan kandidat yang berkontestasi dalam Pemilu Presiden dan Wakil Presiden Indonesia 2024 untuk mengetahui sentimen masyarakat dari video debat kepresidenan. Algoritma yang digunakan untuk klasifikasi sentimen positif dan negatif dilakukan menggunakan algoritma *Support Vector Machine* (SVM) dengan mengimplementasikan metode pelabelan manual dan *lexicon* ke dalam pelabelan sentimennya. Dalam menentukan performa terbaik, model dievaluasi menggunakan *confusion matrix*, *classification report*, dan kurva ROC/AUC. Setelah dilakukan perbandingan dari rasio pembagian data dan juga metode pelabelan sentimen, didapatkan model SVM dengan rasio 90:10 yang menggunakan pelabelan *lexicon* merupakan model yang memiliki performa paling baik. Pada *dataset* Anies – Muhaimin, model menghasilkan nilai akurasi 0,93; nilai *precision* 0,95; nilai *recall* 0,95; nilai *F1-score* 0,95; dan nilai AUC 0,9772. Pada *dataset* Prabowo – Gibran, model menghasilkan nilai akurasi 0,91; nilai *precision* 0,93; nilai *recall* 0,91; nilai *F1-score* 0,92; dan nilai AUC sebesar 0,97. Sementara itu, *dataset* Ganjar – Mahfud dihasilkan nilai akurasi 0,91; nilai *precision* 0,92; nilai *recall* 0,97; nilai *F1-score* 0,94; dan nilai AUC 0,96. Penggunaan metode *lexicon* meningkatkan skor evaluasi sebanyak 1-7% lebih tinggi daripada pelabelan manual di berbagai rasio pembagian data.

Kata kunci—*Analisis Sentimen, Debat Kepresidenan, Lexicon, Support Vector Machine (SVM), YouTube*