

## ABSTRAK

Jumlah data mengalami pertumbuhan yang sangat cepat dalam era sekarang ini. Data dapat berupa *text*, gambar, suara, dan video. Media sosial menjadialah satu faktor pertumbuhan data, setiap orang berekspresi, beropini dan mengeluh di media sosial.

Dalam tugas akhir ini dilakukan analisis sentimen menggunakan dua metode yaitu *Support Vector Machine* dan *Maximum Entropy*. Langkah pertama adalah dilakukan penambangan data menggunakan twitter API dengan *keyword* masing-masing adalah nama calon dalam pilkada DKI. Setelah mengumpulkan data, dilakukan proses *preprocessing*, setelah proses *preprocessing* dilakukan pengambilan fitur pada setiap *tweet*, fitur yang didapatkan kemudian dikumpulkan menjadi sebuah list fitur. List fitur kemudian ditransformasikan menjadi *feature vector* dengan bentuk *binary* kemudian ditransformasikan menggunakan metode *Tf-idf*. Dataset terdiri dari 2 data yaitu *training* dan *testing*. *Training* diberikan label secara manual. Untuk pengujian performa algoritma digunakan metode *K-Fold Cross Validation*.

Hasil pengujian adalah akurasi yang diperoleh mencapai rata-rata 75% dengan komposisi data *training* dan data uji 90:10. *Kernel* yang paling optimal dalam pengujian adalah *kernel* linier. Perubahan jumlah *folding* tidak berpengaruh terhadap tingkat akurasi. Metode *Support Vector Machine* lebih baik digunakan daripada *Maximum Entropy*; untuk melakukan analisis sentimen.

Kata kunci : *Support Vector Machine, Maximum Entropy, Text Mining*