

## ABSTRAK

Platform e-commerce Shopee menghasilkan jutaan ulasan produk setiap harinya yang mengandung opini dan pengalaman pembeli. Besarnya volume ulasan yang terus bertambah setiap harinya membuat pendekatan manual dalam menganalisis opini konsumen menjadi tidak praktis, sehingga dibutuhkan sistem klasifikasi sentimen yang bekerja secara otomatis. Studi ini mengembangkan sistem klasifikasi sentimen berbasis Multinomial Naïve Bayes yang dioptimalkan melalui penerapan seleksi fitur Chi-Square pada data ulasan produk elektronik di platform Shopee. Pengumpulan data dilaksanakan secara otomatis menggunakan teknik web scraping berbasis Python yang memanfaatkan library Selenium dan Undetected ChromeDriver, menghasilkan 2.118 ulasan dari kategori televisi, handphone, laptop, dan headset. Setelah proses pelabelan, pembersihan data noise (data cleaning), dan preprocessing teks mencakup tahapan case folding, cleansing, tokenizing, penghapusan stopword, dan stemming yang diimplementasikan menggunakan library Sastrawi, menghasilkan dataset final berjumlah 1.402 data dengan komposisi 1.214 data berkelas Positif (86,59%) dan 188 data berkelas Negatif (13,41%). Proses vektorisasi TF-IDF menghasilkan 2.678 fitur yang selanjutnya disaring menggunakan Chi-Square hingga tersisa 500 fitur terpilih, setara dengan reduksi dimensi sebesar 81,33%. Pengujian model tanpa seleksi fitur menghasilkan akurasi 90,04%, presisi 91,07%, recall 90,04%, dan F1-Score 87,40%. Setelah penerapan seleksi fitur Chi-Square, performa model meningkat dengan akurasi 91,10%, presisi 90,88%, recall 91,10%, dan F1-Score 89,59% dengan nilai cross-validation sebesar  $89,21\% \pm 0,76\%$ . Hasil penelitian membuktikan bahwa kombinasi ketiga metode tersebut terbukti andal dalam memetakan polaritas sentimen ulasan produk elektronik berbahasa Indonesia pada platform Shopee.

**Kata Kunci:** Analisis Sentimen, Chi-Square, E-Commerce, Naïve Bayes, Shopee, Text Mining, TF-IDF