

Aspect and Opinion Extraction of Indonesian Lipsticks Product Reviews using Conditional Random Field (CRF)

Deffri Kun Indarta¹, Ade Romadhony²

^{1,2} School of Computing, Telkom University, Bandung

¹deffrikunindarta@students.telkomuniversity.ac.id, ²aderomadhony@telkomuniversity.ac.id

Abstrak

Ekstraksi aspek dan opini merupakan kunci utama dalam beberapa aplikasi, seperti market analysis. Ekstraksi aspek dan opini ulasan produk lipstik dalam Bahasa Indonesia di modelkan menggunakan metode Conditional Random Field (CRF). Ulasan produk memiliki informasi penting, terdiri dari aspek dan opini yang sangat berpengaruh bagi pembeli untuk membuat keputusan terhadap produk. Namun, dengan banyaknya ulasan produk yang tersedia di media sosial dan e-commerce, banyak opini dan aspek dapat diperoleh dalam satu ulasan produk, menganalisa serta mengekstraksi aspek dan opini menjadi semakin sulit dan memakan waktu. Dataset yang digunakan berisi teks ulasan yang ditulis tidak hanya dalam bahasa Indonesia standar dan sehari-hari tetapi juga bahasa Inggris standar dan bahasa sehari-hari, dilabeli menggunakan format notasi BIO. Hasil eksperimen menunjukkan bahwa nilai average F1 dari 10 jenis label aspek dan opini adalah 44.1% dengan akurasi 81.8%. Hasil evaluasi model berdasarkan metode HMM menunjukkan nilai F1 dan akurasi yang lebih rendah. Error yang ditemukan pada hasil ekstraksi kebanyakan muncul pada saat kata yang sama memiliki makna yang berbeda. Jenis error ini ditemukan sebanyak 75% dari keseluruhan hasil ekstraksi. Error kedua terbesar disebabkan oleh kata yang tidak diketahui atau baru dengan jumlah 14% dari keseluruhan hasil ekstraksi.

Kata kunci : ulasan produk, ekstraksi aspek, ekstraksi opini, Conditional Random Field.
