

1. Pendahuluan

Latar Belakang

Ulasan merupakan sebuah teks yang berfungsi untuk mengukur, menilai, dan memikirkan kritik mengenai karya atau kejadian yang diulas [1]. Sebuah teks ulasan umumnya mengandung informasi opini terhadap aspek produk. Aspek merupakan kata atau frasa yang menggambarkan atribut atau fitur entitas yang menjadi target opini. Opini merupakan kata atau frasa yang menunjukkan emosi subjektif terhadap atribut atau fitur entitas [3].

Sebuah ulasan produk berguna bagi konsumen untuk menentukan apakah produk tersebut layak untuk dibeli atau tidak. Menurut survei yang dilakukan *Bright Local*, hanya 48% konsumen yang akan melanjutkan transaksi jika penilaian produk kurang dari empat dalam skala lima [2]. Hal ini penting bagi pemilik bisnis karena dengan adanya ulasan produk, mereka mendapatkan wawasan dan informasi terhadap produknya sehingga dapat meningkatkan kualitas produknya. Namun dengan banyaknya ulasan yang terdapat pada situs-situs ulasan seperti *femaledaily.com* membuat pemilik bisnis kesulitan untuk memutuskan aspek apa saja yang sering diulas dalam produknya membutuhkan waktu dan sumber daya yang besar jika dilakukan secara manual. Dengan adanya masalah tersebut, akan dibangun sebuah sistem yang menangani ekstraksi aspek pada ulasan produk lipstik.

Ekstraksi aspek merupakan subugas dari analisis sentimen yang terdiri dari pengidentifikasian target opini dalam teks opini [4]. Terdapat berbagai metode yang telah diterapkan untuk melakukan ekstraksi aspek, salah satunya menggunakan *Pattern Knowledge* dan *Taxonomy* yang menghasilkan *F1 score* sekitar 20%-69%, dan *Recall* sebesar 89,3% [5].

Dalam Tugas Akhir ini, ekstraksi aspek dilakukan dengan metode *Bidirectional Long Short-Term Memory* dengan *Conditional Random Fields* (BiLSTM-CRFs). Dalam penelitian yang dilakukan oleh Athanasios [7], BiLSTM-CRFs memberikan performansi yang tinggi dalam melakukan ekstraksi aspek pada dataset ulasan laptop dan restoran yang menghasilkan *F1 Score* sebesar 88%. BiLSTM dirancang untuk menyandikan setiap elemen dengan mempertimbangkan konteks kiri dan kanan. CRFs memiliki keunggulan dibandingkan metode lainnya seperti *Hidden Markov Model* dalam melakukan segmentasi, dan pelabelan secara terurut [6].

Topik dan Batasannya

Berdasarkan dari latar belakang tersebut, dapat dirumuskan permasalahan yang dapat diangkat dalam tugas akhir ini, yaitu bagaimana membangun dan mengimplementasikan sistem yang dapat mengekstraksi aspek menggunakan metode *Bidirectional Long Short-Term Memory* dengan *Conditional Random Fields* (BiLSTM-CRFs). Adapun Batasan-batasan yang ada dari permasalahan ini adalah

- Dataset yang digunakan merupakan ulasan produk lipstik Bahasa Indonesia dari situs *femaledaily.com*.
- Dataset terdiri dari 1.429 kalimat, lalu ditokenisasi menjadi 23.004 token/kata.
- Pelabelan aspek menggunakan notasi *beginning-inside-outside* (BIO) dan dibagi kedalam enam aspek, yaitu warna, kelembaban, pigmentasi, kemasan, harga, dan ketahanan warna. Tabel 1 menampilkan informasi label setiap aspek.

Tabel 1. Informasi Label

Label	Aspek
B-ASP-COL	Warna
I-ASP-COL	
B-ASP-MOI	Kelembaban
I-ASP-MOI	
B-ASP-PGM	Pigmentasi
I-ASP-PGM	
B-ASP-PKG	Kemasan
I-ASP-PKG	
B-ASP-PRI	Harga
I-ASP-PRI	
B-ASP-STA	Ketahanan warna
I-ASP-STA	
O	Tidak ada aspek

Tujuan

Tujuan dari penelitian ini adalah untuk membangun sebuah sistem yang dapat mengekstraksi aspek menggunakan metode *Bidirectional Long-Short Term Memory* dengan *Conditional Random Fields* (BiLSTM-CRFs) berdasarkan hasil kinerja *F1 Score* dan melakukan analisis terhadap hasil prediksi.

Organisasi Tulisan

Laporan tugas akhir ini terdiri dari lima bagian. Pada bagian pertama menjelaskan pendahuluan, pada bagian kedua menjelaskan studi terkait, pada bagian ketiga menjelaskan sistem yang dibangun, pada bagian keempat menjelaskan evaluasi dari sistem yang dibangun, dan pada bagian kelima menjelaskan kesimpulan yang diambil dari tugas akhir ini.