

Abstrak

Media sosial adalah forum di mana individu berkumpul dan bertukar informasi, sehingga memfasilitasi komunikasi kontemporer. Konten media sosial menghadirkan tantangan unik untuk ekstraksi informasi karena sifatnya yang beragam dan tidak terstruktur. Misalnya, konten yang paling umum di platform media sosial seperti Twitter/X bersifat komersial, biasanya terdiri dari iklan produk atau promosi produk. Hambatan utama untuk menerima informasi adalah struktur gaya penulisan pengguna yang beragam dan bervariasi, yang menyajikan berbagai format. Banyaknya format tulisan dalam informasi yang dimaksud membuatnya kurang efisien. Di sinilah peran ekstraksi informasi masuk, mengubah informasi tidak terstruktur menjadi format terstruktur menggunakan metode *Bidirectional Long-Short Term Memory with Conditional Random Fields* (BiLSTM-CRF). Metode ini dipilih karena memberikan konteks untuk informasi dari masa lalu dan masa depan, yang cocok untuk tugas mengekstraksi informasi. Tujuan dari penelitian ini adalah untuk mengekstrak informasi dengan menggunakan *Bidirectional Long Short-Term Memory* (BiLSTM) dengan *Conditional Random Fields* (CRF) untuk klasifikasi. Proses ini melibatkan ekstraksi informasi sesuai dengan label BIO, yang kemudian disusun dan dapat diakses untuk analisis. Hasil yang diperoleh dari penerapan model yang disebutkan di atas menghasilkan nilai kinerja berikut: *precision* 87%, *recall* 77%, dan *f1-score* 80%.

Keywords-social media, product advertisements, BiLSTM CRF, information extraction