

1. Pendahuluan

Latar Belakang

Dalam era informasi digital saat ini, artikel berita menjadi salah satu sumber utama bagi masyarakat untuk memperoleh informasi terkini. Kualitas penulisan artikel berita sangat mempengaruhi pemahaman dan kepercayaan pembaca terhadap informasi yang disampaikan. Namun, sering kali ditemukan kesalahan penulisan seperti kesalahan ejaan, tata bahasa, dan penempatan tanda baca yang dapat mengurangi kredibilitas dan keakuratan berita [1], [2].

Kesalahan penulisan pada artikel berita dapat disebabkan oleh berbagai faktor, termasuk kesalahan manusia saat mengetik, kurangnya waktu untuk proses penyuntingan, serta tekanan untuk menyajikan berita secepat mungkin [3], [4]. Kesalahan-kesalahan ini tidak hanya mempengaruhi kualitas berita tetapi juga dapat menyebabkan salah tafsir informasi oleh pembaca [5].

Pentingnya kualitas penulisan yang baik telah mendorong banyak penelitian dalam bidang koreksi teks otomatis. Salah satu solusi yang menjanjikan adalah penggunaan model bahasa berbasis pembelajaran mesin yang mampu mendeteksi dan mengoreksi kesalahan penulisan secara otomatis [6]. Model bahasa ini dilatih menggunakan dataset yang berisi berbagai contoh kesalahan penulisan, sehingga mampu mempelajari pola-pola kesalahan yang sering terjadi dan memberikan koreksi yang tepat [7].

Penggunaan model bahasa untuk koreksi teks pada artikel berita tidak hanya bertujuan untuk memperbaiki kesalahan penulisan, tetapi juga untuk meningkatkan efisiensi proses penulisan dan penyuntingan [8]. Dengan adanya teknologi ini, jurnalis dan editor dapat lebih fokus pada konten berita itu sendiri tanpa harus khawatir tentang kesalahan penulisan yang dapat mengurangi kualitas artikel [9].

Penelitian ini bertujuan untuk mengimplementasi sistem koreksi teks menggunakan model bahasa yang mampu mendeteksi dan mengoreksi kesalahan penulisan pada artikel berita dengan tingkat akurasi yang cukup tinggi [10]. Dengan demikian, diharapkan dapat memberikan kontribusi dalam meningkatkan kualitas penulisan artikel berita dan memastikan informasi yang disampaikan kepada masyarakat lebih akurat dan mudah dipahami [11].

Topik dan Batasannya

Penelitian ini berfokus pada analisis performa dari sistem koreksi teks menggunakan model bahasa untuk mendeteksi dan mengoreksi kesalahan penulisan pada artikel berita bahasa Indonesia. Sistem koreksi teks yang digunakan dalam penelitian ini diimplementasikan dengan menggunakan model bahasa IndoBERT yang mampu mengidentifikasi kesalahan penulisan, serta memberikan koreksi yang sesuai [12]. Dalam implementasi koreksi teks menggunakan model bahasa, penelitian ini menggunakan ruang lingkup artikel berita berbahasa Indonesia. Model bahasa dilatih dan divalidasi menggunakan dataset yang terdiri dari artikel berita yang mengandung berbagai jenis kesalahan penulisan yang sering ditemukan dalam praktik jurnalistik [13].

Tujuan

Tujuan dari penelitian ini adalah untuk mengimplementasi sistem koreksi teks dengan menggunakan model bahasa yang mampu mendeteksi dan mengoreksi kesalahan penulisan pada artikel berita berbahasa Indonesia dan juga menganalisis performa dari sistem koreksi teks tersebut dengan menggunakan metrik evaluasi yang relevan. Manfaat yang diharapkan dari implementasi sistem koreksi teks ini adalah untuk meningkatkan kualitas penulisan artikel berita bahasa Indonesia, memastikan akurasi informasi, dan memberikan kontribusi dalam pengembangan teknologi koreksi teks otomatis di Indonesia.