

BAB I

PENDAHULUAN

1.1 Latar Belakang

Peristiwa merujuk pada kejadian atau rangkaian kejadian bermakna yang terjadi dalam berbagai konteks, seperti sejarah, sastra, atau kehidupan sehari-hari. Secara umum, peristiwa mencakup segala hal yang terjadi, baik biasa maupun luar biasa, yang sering kali berdampak signifikan pada kehidupan masyarakat. Peristiwa-peristiwa luar biasa, seperti bencana alam, konflik sosial, atau perubahan politik, sering menarik perhatian publik karena pengaruhnya yang luas, baik di tingkat lokal maupun internasional. Dalam konteks ini, kebutuhan akan informasi yang cepat dan akurat menjadi semakin penting, terutama bagi masyarakat yang bergantung pada berita *online* untuk mengikuti perkembangan terbaru (Theofany et al., 2024).

Berita online memiliki peran penting dalam menyebarkan informasi mengenai peristiwa-peristiwa yang terjadi, baik di tingkat lokal maupun global. Melalui berita *online*, masyarakat dapat memahami situasi terkini, mengambil langkah yang tepat dalam merespons suatu kejadian, dan bahkan membentuk opini publik. Berita online memberikan kecepatan akses informasi dan jangkauan yang lebih luas dibandingkan media cetak atau televisi konvensional, sehingga menjadi sumber utama informasi bagi banyak orang. Berita yang cepat dan akurat berfungsi sebagai sarana edukasi, membantu masyarakat untuk lebih sadar akan kejadian yang mungkin berdampak langsung pada mereka. Meskipun berita *online* berperan penting dalam menyebarkan informasi, tidak semua daerah mendapatkan akses yang memadai terhadap berita ini, terutama di daerah terpencil atau yang kurang mendapat perhatian media arus utama.

Akses informasi di wilayah pelosok seperti Kabupaten Simeulue, Provinsi Aceh, masih terbatas. Kabupaten ini sering kali mengalami peristiwa penting, termasuk bencana alam dan isu sosial, tetapi kurang mendapat perhatian dari media nasional. Masyarakat setempat terpaksa bergantung pada sumber informasi lokal untuk mengetahui perkembangan di daerah mereka, meskipun banyak di antara

peristiwa tersebut memiliki dampak yang signifikan. Keterbatasan akses ini menunjukkan adanya kebutuhan akan metode klasifikasi dan ekstraksi informasi yang mampu menyediakan akses berita yang relevan dengan lebih cepat dan tepat.

Salah satu pendekatan yang dapat mengatasi permasalahan ini adalah *Named Entity Recognition* (NER), sebuah teknik dalam *Text Mining* dan *Natural Language Processing* (NLP) yang bertujuan untuk menemukan dan mengkategorikan entitas dalam teks seperti nama, lokasi, dan organisasi. (Azizi et al., 2023). Melalui penerapan NER, informasi penting dapat lebih mudah diakses dan disaring sehingga berita signifikan di wilayah seperti Simeulue dapat tersebar dengan lebih efisien (Rifani et al., 2019).

Seiring dengan perkembangan teknologi, metode NER terus mengalami peningkatan, terutama dengan dukungan algoritma *machine learning* dan *deep learning*. Beberapa model seperti *Long Short-Term Memory* (LSTM), *Hidden Markov Model* (HMM), dan *Conditional Random Field* (CRF) telah menunjukkan efektivitasnya dalam meningkatkan akurasi ekstraksi entitas (Zahra, 2021). Dalam konteks berita Simeulue, penerapan model *deep learning* yang relevan diharapkan mampu memberikan hasil klasifikasi yang lebih akurat. (Batbaatar & Ryu, 2019).

Teknologi BiLSTM (*Bidirectional Long Short-Term Memory*) dan CNN (*Convolutional Neural Networks*) adalah kombinasi yang efektif untuk deteksi entitas dalam teks. BiLSTM mampu menangkap konteks kalimat secara dua arah, sedangkan CNN dapat mengekstraksi fitur lokal dalam pola kata (Lin & Liu, 2022). Kombinasi ini memungkinkan model mengenali entitas peristiwa dengan lebih akurat. Penelitian (Sukardi et al., 2021) menunjukkan bahwa model BiLSTM-CNNs yang diterapkan pada teks berbahasa Indonesia dengan *embedding word2vec* berhasil mencapai skor F1 sebesar 71.37%, yang mencerminkan efektivitas model ini dalam mendeteksi entitas dari berita formal maupun informal.

Berdasarkan fenomena tersebut, penelitian ini berfokus pada pengembangan model *Named Entity Recognition* (NER) untuk mendeteksi entitas peristiwa dalam berita *online* Kabupaten Simeulue, yang memiliki variasi bahasa formal dan informal. Dengan menerapkan pendekatan *deep learning*, khususnya menggunakan model BiLSTM-CNNs, diharapkan model ini dapat secara efektif menangani

variasi bahasa tersebut sehingga informasi yang relevan dapat diakses lebih cepat dan akurat oleh masyarakat Simeulue.

1.2 Rumusan Masalah

Berdasarkan pemaparan latar belakang di atas, masalah utama yang dibahas dalam penelitian ini dapat dirumuskan sebagai berikut:

1. Bagaimana membangun sebuah sistem berbasis web yang dapat mendeteksi entitas peristiwa dalam berita *online* dengan menggunakan model *Named Entity Recognition* (NER) berbasis BiLSTM dan CNNs?
2. Bagaimana mengukur performa gabungan metode BiLSTM dan CNNs dalam mendeteksi entitas peristiwa kejadian pada suatu berita *online*?

1.3 Tujuan Penelitian

Penelitian ini bertujuan untuk:

1. Membangun sebuah sistem berbasis web yang mampu melakukan *Named Entity Recognition* (NER) pada berita *online* untuk mendeteksi entitas peristiwa, lokasi, dan waktu.
2. Mengembangkan model *Named Entity Recognition* (NER) khusus untuk domain peristiwa kejadian dengan memanfaatkan kombinasi metode BiLSTM dan CNNs.
3. Mengevaluasi kinerja kombinasi BiLSTM dan CNNs dalam mendeteksi entitas peristiwa pada pemberitaan, dengan mengukur *F1-Score* yang dihasilkan

1.4 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini meliputi:

1. Memberikan solusi berbasis web yang dapat membantu pengguna dalam mengidentifikasi entitas peristiwa, lokasi, dan waktu dalam berita *online* secara otomatis.

2. Menghasilkan model Named Entity Recognition yang dapat meningkatkan pemahaman terhadap pola entitas dalam berita lokal, khususnya di Kabupaten Simeulue
3. Mendukung pengembangan teknologi *Natural Language Processing* (NLP) untuk ekstraksi informasi dari teks berita, sehingga dapat digunakan untuk berbagai kebutuhan seperti analisis tren berita atau sistem rekomendasi informasi.

1.5 Ruang Lingkup dan Batasan Penelitian

Ruang lingkup dan batasan penelitian ini ditetapkan untuk memperjelas cakupan studi serta menghindari perluasan pembahasan yang tidak relevan. Penelitian ini berfokus pada penerapan model *Named Entity Recognition* (NER) menggunakan arsitektur BiLSTM-CNNs dalam mengidentifikasi entitas pada berita *online*. Adapun ruang lingkup dan batasannya adalah sebagai berikut:

1. Data yang digunakan hanya artikel berita mengenai peristiwa yang terjadi di Kabupaten Simeulue sejak tahun 2019 hingga 2024.
2. Informasi yang diidentifikasi hanya lokasi, peristiwa dan tanggal terjadinya peristiwa.
3. Penelitian ini hanya menggunakan bahasa Indonesia sebagai bahasa utama dalam teks berita, sehingga tidak mencakup artikel dalam bahasa asing atau teks campuran.
4. Label entitas yang digunakan terbatas pada tiga jenis yaitu *Event*, *Loc*, dan *Time*, serta label O untuk token di luar entitas, sehingga entitas lain seperti nama orang atau organisasi tidak dipertimbangkan.
5. Penelitian ini tidak mencakup analisis sentimen, klasifikasi topik, atau fitur linguistik lainnya, melainkan hanya fokus pada ekstraksi entitas bernama sesuai kategori yang telah ditentukan.