

1. Pendahuluan

Latar Belakang

Analisis sentimen pada media sosial saat ini menjadi suatu topik menarik untuk diteliti, hal ini dikarenakan perkembangan penggunaan media sosial yang pesat terutama di Indonesia[1]. Salah satunya adalah media sosial yang bernama Twitter. Dengan mengirimkan cuitan yang merupakan pesan singkat, pengguna Twitter dapat menyampaikan pandangannya terhadap suatu bahasan tertentu yang terjadi di sekitarnya. Cuitan ini dapat dianalisis sentimennya untuk berbagai kebutuhan, misalnya untuk melihat kepribadian seorang individu atau ketertarikan orang terhadap suatu hal dan banyak hal lainnya [2][3].

Bencana pada beberapa tahun terakhir banyak terjadi di mana-mana, seperti banjir, gempa bumi, gunung meletus. Topik tentang bencana ini tidak luput dari perhatian warganet untuk membahas atau membagikan pandangannya terhadap bencana yang terjadi di Indonesia. Warganet biasanya membicarakan banyak hal tentang topik bencana ini, mulai dari penyebab bencana dan bagaimana penanganan dari pihak yang berwenang dalam menangani bencana yang terjadi. Cuitan mengenai bencana ini dapat dianalisis sentimennya menggunakan sistem analisa sentimen dan bisa menjadi rekomendasi kepada pihak yang berwenang untuk mengambil tindakan atau melihat penilaian warganet terhadap kinerjanya.

Untuk membangun sistem analisa sentimen ini akan digunakan metode *word embedding*. *Word embedding* adalah metode yang berguna untuk merepresentasikan kata ke dalam bentuk vektor. Metode ini dapat meningkatkan performa analisa sentimen, oleh karena hal tersebut tersebut *word embedding* banyak digunakan pada penelitian yang membahas analisa sentimen. Dalam penelitian ini akan digunakan model *word embedding* yang bernama *GloVe (Global Vector)*. Model *Global Vector* dipilih karena memiliki tingkat akurasi yang baik dibandingkan dengan model *word embedding* lainnya seperti *word2vec (Continuous Bag of Words & Skip-gram)* dan *doc2vec*[4].

Word embedding dan data cuitan kemudian akan dianalisis untuk diklasifikasikan sentimennya. Pada penelitian ini akan digunakan metode *deep learning* untuk mengklasifikasikan sentimen. Model yang akan digunakan adalah *Long Short-Term Memory (LSTM)*. *LSTM* ini memiliki beberapa layer yang salah satu layernya berfungsi untuk *word embedding* dan memiliki performa yang baik untuk mengklasifikasikan sentimen, jika digunakan dengan *word embedding* model *Global Vector* [4].

Pada penelitian ini, data yang digunakan adalah data cuitan yang bertemakan bencana dengan jumlah label data seimbang dan dengan label data sentimen tidak seimbang. Data dari Wikipedia Indonesia digunakan sebagai korpus, karena belum adanya model *pre-trained Global Vector* dalam Bahasa Indonesia. Data dari Wikipedia Indonesia akan dibentuk menjadi model *Global Vector* yang digunakan sebagai *word embedding*. Penjelasan lebih lengkap mengenai pelatihan ini akan dijelaskan pada bagian selanjutnya.

Topik dan Batasannya

Topik dan batasan pada penelitian adalah menganalisa cuitan tentang bencana pada di media sosial Twitter berbahasa Indonesia. Jumlah data cuitan yang digunakan pada penelitian ini berjumlah 514, 1150 dan 1506 data cuitan, yang mana telah dilabeli dengan label positif dan negatif dengan jumlah yang sama. Selain itu, pada penelitian ini akan digunakan data dengan label yang tidak seimbang dengan total 1282 data cuitan. Dengan 751 data label positif dan 531 data label negatif. Model *word embedding* yang digunakan adalah model yang dibentuk dari kumpulan data Wikipedia Indonesia dengan 364998 kosakata, dikarenakan belum ada model *Global Vector* Bahasa Indonesia yang tersedia. Model *Global Vector* yang dibentuk menggunakan dimensi 50, 100, 150, 200, 250, 300 dan 350.

Permasalahan

Bagaimana penerapan *word embedding Global Vector* dan algoritma *deep learning Long Short-Term Memory* untuk membentuk model analisa sentimen.

Tujuan

Menerapkan dan menganalisa performa model analisa sentimen yang menggunakan *word embedding Global Vector* dan model *deep learning Long Short-Term Memory* untuk mengklasifikasikan sentimen cuitan bencana.