

ABSTRACT

In the current era, the development of the internet is increasing very fast, a lot of information that we can easily get through the internet. One area affected by the rapid development of the internet is the area of hotel accommodation services, many people more often visit the site to find information on lodging near their destination for travel, whether for travel, business trips or just wanting to take a vacation. One of the biggest hotel accommodation service sites in Indonesia is Traveloka. The site provides hotel services and transportation ticket services. With a large number of visitors using Traveloka's services to search for hotel services, so many reviews were written about the hotel they visited. The hotel reviews can be used to build a sentiment analysis system that can provide advice for other visitors in finding hotel services that suit their needs.

This study aims to build sentiment analysis using the Long Short-Term Memory (LSTM) method. The data used in this study came from the Traveloka website, the data will be classified into two sentiment polarities, namely positive and negative. Researchers also do tuning in LSTM parameters, such as dropouts, and batch sizes. In addition, researchers also used two different pre-training, namely word2vec and glove to determine the effect of these parameters on the accuracy of performance in the LSTM classification. Furthermore, the results of this study will be compared with previous studies.

Twelve times the experiment was carried out using a combination of parameters and word embedding to create a different LSTM model. The results of this study indicate that tuning the Long Short-Term Memory method can produce an accuracy value of 96.3%, better than previous studies using the Naive Bayes method combined with the lexicon base which obtained an accuracy of 89.95%. So the current research can improve the accuracy of 6.35%. With these results, the research conducted can make the sentiment analysis system more accurate and can help others who want to find hotel accommodations.

Keywords : Sentiment analysis, LSTM, Dropout, Batch Size, Pre-training

INTISARI

Pada era saat ini perkembangan internet meningkat sangat cepat, banyak informasi yang bisa kita dapat dengan mudah melalui internet. Salah satu bidang yang terkena dampak dari pesatnya perkembangan internet adalah bidang layanan akomodasi hotel, banyak orang lebih sering mengunjungi situs untuk mencari informasi tempat penginapan yang dekat dengan tempat tujuan mereka bepergian, baik untuk wisata, perjalanan bisnis maupun sekedar ingin berlibur. Salah satu situs layanan akomodasi hotel terbesar di Indonesia adalah Traveloka. Situs tersebut menyediakan layanan hotel maupun tiket layanan transportasi. Dengan banyaknya pengunjung yang menggunakan jasa Traveloka untuk mencari layanan hotel, maka banyak pula ulasan yang ditulis tentang hotel yang mereka singgahi. Ulasan hotel tersebut dapat dimanfaatkan untuk membangun sistem analisis sentimen yang dapat memberikan saran bagi para pengunjung lain dalam mencari layanan hotel yang sesuai dengan kebutuhan mereka.

Penelitian ini bertujuan untuk membangun analisis sentimen dengan menggunakan metode *Long Short-Term Memory* (LSTM). Data yang digunakan pada penelitian ini berasal dari situs Traveloka, data tersebut akan diklasifikasikan menjadi dua polaritas sentimen, yaitu positif dan negatif. Peneliti juga melakukan *tuning* pada parameter LSTM, seperti *dropout*, dan *batch size*. Selain itu peneliti juga menggunakan dua *pre-training* yang berbeda yaitu *word2vec* dan *glove* untuk mengetahui pengaruh parameter tersebut terhadap performa akurasi pada klasifikasi LSTM. Selanjutnya hasil dari penelitian ini akan dibandingkan dengan penelitian sebelumnya.

Ekperimen dilakukan sebanyak dua belas kali dengan menggunakan kombinasi antara parameter dan *word embedding* untuk membuat model LSTM yang berbeda. Hasil dari penelitian ini menunjukkan bahwa melakukan *tuning* pada metode *Long Short-Term Memory* dapat menghasilkan nilai akurasi yaitu 96.3%, lebih baik dibandingkan penelitian sebelumnya yang menggunakan metode *naive bayes* yang dikombinasikan dengan *lexicon base* yang memperoleh akurasi sebesar 89.95%.

Sehingga penelitian yang dilakukan saat ini dapat meningkatkan akurasi sebesar 6.35%. Dengan hasil tersebut maka penelitian yang dilakukan dapat menjadikan sistem analisis sentimen menjadi lebih akurat dan dapat membantu orang lain yang ingin mencari akomodasi hotel.

Kata kunci – Analisis Sentimen, LSTM, *Dropout*, *Batch Size*, *Pre-training*