

INTISARI

PENENTUAN DESTINASI WISATA FAVORIT BERBASIS ATURAN DAN ANALISIS SENTIMEN PADA TWEET BERBAHASA INDONESIA

Oleh

NUR RISTA KURNIAWATI

12/334855/PA/15044

Peningkatan industri pariwisata Indonesia saat ini tidak lepas dari peran teknologi. Selama berwisata turis juga akan membagikan pengalamannya secara *real-time* melalui *upload* foto maupun video serta aktif berkomentar di media sosial. Salah satu media sosial yang digemari di Indonesia yaitu Twitter. Dari tweet dapat diketahui destinasi wisata yang sedang hangat dibicarakan oleh masyarakat maupun wisatawan yang berkunjung di suatu kota. Selain itu tweet tersebut dapat mengandung informasi seputar kondisi pariwisata, pengalaman pengunjung, opini pengunjung di suatu tempat wisata, dan rekomendasi tempat wisata lainnya. Untuk itu diperlukan suatu sistem untuk memberikan informasi seputar tren pariwisata di suatu kota. Tentu saja informasi ini sangat bermanfaat untuk semua kalangan mulai dari wisatawan, masyarakat, pemerintah, hingga industri pariwisata di kota tersebut.

Dalam penelitian ini dibuat sistem yang dapat menentukan destinasi wisata favorit di suatu kota, memberikan informasi tren masyarakat berkaitan dengan pariwisata melalui *hashtag* favorit, memberikan informasi berupa opini masyarakat maupun wisatawan terhadap pariwisata melalui analisis sentimen, serta memberikan informasi mengenai waktu favorit pengguna Twitter di kota tersebut. Dalam penentuan destinasi wisata favorit digunakan kamus lokasi (*gazetteer*) untuk mengetahui banyaknya tweet di suatu destinasi wisata dan dibantu dengan analisis sentimen pada tweet tersebut. Salah satu pendekatan analisis sentimen adalah menggunakan leksikon sentimen. Pada penelitian ini dikembangkan leksikon sentimen untuk pariwisata berbasis aturan dengan pendekatan kamus.

Dalam pengujian performa sistem dalam mengenali tweet pariwisata dilakukan tiga kali percobaan dengan tiga *search query* berbeda yaitu ‘#Yogyakarta’, ‘#Bali’, dan ‘#Jakarta’. Hasilnya didapat rata-rata *recall* 71,3%, *precision* 99,2%, *f-measure* 82,6%, dan akurasi 93,8%. Sementara itu dalam pengujian performa analisis sentimen dilakukan tiga kali percobaan dengan tiga *search query* berbeda yaitu ‘Yogyakarta’, ‘Bali’, dan ‘Jakarta’ dengan data uji berbeda dengan pengujian sistem dalam mengenali tweet pariwisata atau bukan. Hasilnya didapat rata-rata *recall* 73,1%, *precision* 100%, *f-measure* 84,1%, dan akurasi 93,2%.

Kata kunci: ekstraksi informasi, *social media*, *tourism*, analisis sentimen, leksikon sentimen, *rule-based*, *dictionary-based*, *gazetteer*

ABSTRACT

DETERMINATION TOP TOURIST DESTINATION BASED ON RULE-BASED APPROACH AND SENTIMENT ANALYSIS IN INDONESIAN TWEET

By

NUR RISTA KURNIAWATI

12/334855/PA/15044

Nowadays, improvement of Indonesia's tourism industry can not be separated from technology. During vacation, tourists will share their experiences in real-time via upload photos and videos as well as active comments on social media. One of the most popular social media in Indonesia is Twitter. From tweets will be known top tourist destination that is popular and being discussed by people around city. Beside that, tweet also contain valuable information about the condition of tourist site, tourist's opinion at the tourist sites, and alot of recommendation based on their experiences during the trip. Therefore we need a system to provide information about tourism trends in a city. Of course that information is very useful to all people start from tourists, the community, the government, also the tourism industry in the city.

In this research, system that can determine favorite tourist destination in the city, providing information about trends tourism through the top *hashtag*, provide information such as public opinion in tourism through sentiment analysis, as well as providing information about Twitter user's favorite time has been built. In the determination of top tourist destinations used location dictionary (gazetteer) to find out how many tweets in a tourist sites and then used sentiment analysis on that tweet to know about the opinion of that tourist sites. One of the many approaches to do sentiment analysis is using sentiment lexicon. In this research, sentiment lexicon for tourism based on rule-based with dictionary approach has been developed.

Based on system testing that has been done, the system's performance in recognizing tourism tweet conducted three experiments with three different search query are '#Yogyakarta', '#Bali', '#Jakarta'. The results obtained with average *recall* 71,3%, *precision* 99,2%, *f-measure* 82,6%, and accuracy 93,8%. Then the system's performance in sentiment analysis has been done too with different data testing conducted three experiments with Yogyakarta, Bali, and Jakarta. The results obtained with average *recall* 73,1%, *precision* 100%, *f-measure* 84,1%, and accuracy 93,2%.

Keywords: information extraction, social media, tourism, sentiment analysis, sentiment lexicon, rule-based, dictionary-based, gazetteer