ABSTRACT

With 3 million users and stable customer growth, BOLT! is a service with a high potential to be extracted information using Data Mining approach. By the end of February 2018, BOLT!'s Twitter account reached more than 54,000 followers, 220,000 tweet and replies, and no fewer than 100 complaints per day. In social media, content analysis and text mining are often used to analyze user-generated text and support decision-making. It is an opportunity for the government or even company, to obtain information that is important to improve public satisfaction and enhance the performance of their products and services.

This study proposed the Real-time Company decision support system that handles the complaint from the user by using microblogs data. Data collection in this application is using real-time tweet data which fetched from Streaming API services that provided by Twitter. After that, the raw data tweet being processed using text mining techniques, including preprocessing, the classifying process and extracting location methods. To detect and classify the complaint from the user, this research applied Naïve Bayes classifier. With a total of 15808 tweets data are trained which is balance in the number annotated as a complaint or either non-complaint. Afterward, the model is evaluated by using K-Fold Cross Validation. The result shows the classifier gets the value of Accuracy 94.86%, High precision with 99.18%, and Recall 91.29%. To extract the location from a complaint tweet, NLP technique being used with accuracy 76.8% which is pretty good to predict and get the location. This Real-time Company Decision Support System using Microblogs is very useful for BOLT! in order to support the decision making based on the data retrieved from the system. With Real-time Data Streaming, Real-time Map visualization, and data analytics tools, this decision support system will be needed for BOLT! or other companies that used microblogs as media for customer service.

Keywords*: Twitter, Text Mining, Naïve Bayes, Microblog, Location Extraction*