

INTISARI

Teknik Realtime Data Processing Untuk Monitoring Tweet Menggunakan Analisis Sentimen Dengan Algoritma Naive Bayes

Studi Kasus Tweet Yang Terkait Institusi Ugm

Oleh

Teguh Puji Widiyanto
10/310932/PA/13540

Twitter.com mempunyai lebih dari 316 juta pengguna aktif setiap bulannya dan mempunyai lebih dari 500 Juta *tweet* dipublikasikan setiap hari. Khususnya Indonesia, negara dengan penduduk kurang lebih 255 juta pada tahun 2015 merupakan salah satu negara teraktif dalam menggunakan *Twitter* sebagai media sosial. Karena pengguna *Twitter* di Indonesia lebih dari 50 juta pengguna, menjadikan *Twitter* sangat berpengaruh terhadap kehidupan masyarakat Indonesia. Hal ini dibuktikan dengan beberapa kasus yang menjadi isu nasional karena banyak perhatian di media *Twitter*. Oleh karena itu, diperlukan perhatian khusus untuk meminimalisir dampak dari *Twitter* salah satunya dengan memonitor *tweet* yang relevan secara *realtime*.

Untuk mengatasi kendala tersebut, penelitian ini bertujuan untuk membuat tool monitoring *tweet* menggunakan teknik *data processing* secara *realtime*. Dengan menggunakan analisis sentimen data hasil pengolahan diklasifikasikan dan ditampilkan dalam media monitoring secara *realtime*. Pemrosesan data secara *realtime* dimulai dengan melakukan filter *tweet* dengan kata kunci tertentu sehingga mendapatkan *tweet* yang mendekati atau berkaitan dengan UGM. Hasil filter *tweet* kemudian diproses menggunakan library *Storm* dan diklasifikasikan menjadi *tweet* dengan sentimen positif, negatif dan netral. Analisis sentimen dilakukan dengan menggunakan algoritma *Naive Bayes* dengan terlebih dahulu membuat data latih yang disimpan dalam *Redis Database*.

Library seperti *Twitter4j*, *Apache Storm* dan *Redis* dapat digunakan untuk memproses data secara *realtime*. Teknik tersebut dapat diterapkan untuk memonitoring sentimen *tweet* yang berkaitan dengan Institusi UGM secara *realtime*. Protokol *websocket* dan *NodeJs* membantu untuk menampilkan data hasil prosesing secara *realtime* dengan *latency* yang rendah.

Kata kunci: analisis sentimen, *realtime processing*, monitoring kehumasan

ABSTRACT

Realtime Data Processing Technique To Monitor Tweets Using Sentiment Analysis With *Naive Bayes* Algorithm A Case Study Of Tweets That Related To Ugm

By

Teguh Puji Widiyanto
10/310932/PA/13540

Twitter.com has approximately 316 million active users montly and more than 500 million tweets are published every day. More specific in Indonesia, Country that have aproximately 255 Million citizen is country that has most active user in the world to use Twitter. Indonesia is expected more than 50 Million have Twitter account, so Twitter is most influence social media in indonesian people. Many cases become national effect because published in Twitter. Therefore, It is required spesific attention to minimize impact from using Twitter. One of way to solve the problem is monitor the indonesian tweet that relevan with institution in realtime.

To overcome these obstacles, the study aims to create monitoring tools that use technique of realtime data processing. The result of data processing tweets were classified with Sentiment analisys before tweets are visualized in web browser realltimely. The data processing is started when application filter tweets with specific keyword to get tweets that related with UGM institution. Storm library is used to process filtered tweet in realtime. The result of processing tweet is classified into tweets containing the positive sentiment, negative sentiment and neutral sentiment with Naive Bayes Algorithm. Before it used Naive Bayes algorithm, application need training data that save in Redis Database.

Twitter4J, Apache Storm and Redis Database can be used to process data tweet in realtime. make data processing tool that can monitor tweet in realtime. Websocket protokol and NodeJs can visualize result of processing data in realtime with lower latency.

Keywords: sentimen analysis, *realtime* processing, public relation monitoring