

DETEKSI AKTIVITAS DENGAN MEMANFAATKAN DATA LAYANAN JEJARING SOSIAL BERBASIS LOKASI TWITTER

ABSTRAK

Twitter merupakan salah satu layanan jejaring sosial berbasis lokasi yang kini sangat populer di kalangan masyarakat pengguna internet. Twitter menjadi sumber data yang sangat bermanfaat sekaligus menjadi salah satu pusat penyedia informasi yang bersifat *real-time*. Pada jurnal ini penulis menitikberatkan penelitian pada pendeteksian aktivitas dengan memanfaatkan kumpulan *tweet* para pengguna Twitter. Untuk melakukan penelitian ini digunakan metode-metode seperti metode klasifikasi dengan algoritma Naive Bayes dan metode *clustering* dengan algoritma K-Means. Metode klasifikasi digunakan untuk memisahkan kumpulan *tweet* ke dalam kelas-kelas yang telah ditentukan, dan metode *clustering* digunakan untuk mengelompokkan kumpulan *tweet* yang telah diklasifikasi ke dalam *cluster*-nya masing-masing berdasarkan informasi aktivitas yang terdapat di dalamnya, dan setiap *cluster* mewakili satu aktivitas. Deteksi aktivitas dengan metode klasifikasi dan *clustering* memberikan hasil yang terbilang baik, dibuktikan dengan nilai F-Measure yang diperoleh untuk metode klasifikasi dengan algoritma Naive Bayes yaitu sebesar 77,068 %, dan nilai *purity* untuk metode *clustering* dengan algoritma K-Means yaitu sebesar 0,599.

Kata Kunci: Aktivitas, Twitter, *Tweet*, Klasifikasi, *Clustering*, Naive Bayes, K-Means, *Purity*.

ACTIVITY DETECTION USING DATA OF LOCATION-BASED SOCIAL NETWORK SERVICE TWITTER

ABSTRACT

Twitter is a location-based social networking service that is now very popular among netizens. Twitter is a very useful source of data and become one of the central provider of real time information. In this paper the authors study focuses on detecting activity by utilizing a collection of Twitter users tweet. To conduct this study, used methods such as the method of classification with Naive Bayes algorithm and method of clustering with K-Means algorithm. Classification method is used to separate the tweet into classes that have been determined, and the clustering method is used to classify tweets that have been classified into each cluster based on the information about the activities contained in it. Each cluster represents a single event. Activity detection with classification and clustering methods give fairly good results, evidenced by the value obtained for the F-Measure of the naive Bayes classification method of 77.068%, and the value of purity for clusering method with K-Means algorithm by 0.599.

Keywords: *Activity, Twitter, Tweet, Classification, Clustering, Naive Bayes, K-Means, Purity.*