

Intisari

Data yang dihasilkan dari jejaring sosial seperti Twitter akan menjadi sangat berguna apabila kita dapat mengkomunikasikannya kepada pengguna dengan baik. Salah satu caranya adalah dengan melakukan visualisasi statistik data Twitter.

Tujuan dari penelitian ini adalah membuat sistem klasifikasi topik dengan algoritme *Naive Bayes*, dan visualisasi statistik data Twitter yang berasal dari negara Indonesia menggunakan *tools* Elasticsearch sebagai media penyimpanan dan Kibana sebagai alat *dashboard* visualisasi data pada Elasticsearch.

Dari hasil uji coba terhadap 300 tweet, akurasi yang didapat adalah 84%. Berdasarkan hasil visualisasi pengujian yang dilakukan secara *realtime*, dari 2.508 tweet, 56% adalah “Sosial”, 30% adalah “tidak tahu”, 5,5% adalah “Hiburan”, 3,4% adalah “Olahraga dan Kesehatan”, 2,7% adalah “Berita” dan 2,1% adalah “Teknologi”.

Kata kunci : Klasifikasi, Naive Bayes, Visualisasi, Twitter, Elasticsearch.

Abstract

Data from social networks like Twitter would be very useful if we can communicate it well to user. Visualizing the statistical of Twitter data is one way to do it.

The purpose of this research is to create a topic classification system with Naive Bayes algorithm, and visualization of data statistics Twitter from Indonesia using Elasticsearch as a storage device and Kibana as a dashboard of data visualization on Elasticsearch.

The testing result to 300 tweets, the accuracy is 84 %. Based on the results of tests visualization in real-time, from 2,508 tweets, 56 % are "Social", 30 % are "Don't know", 5.5 % are "Entertainment", 3.4 % are "Sports and Health", 2, 7 % is "News" and 2.1 % are "Technology".

Keywords : Classification, Naive Bayes, Visualization, Twitter, Elasticsearch.