

Pendeteksian topik berita pada twitter dengan menggunakan nonnegative matrix factorization berbasis metode langsung = Topic detection of news on twitter using nonnegative matrix factorization based direct method

Ibtisami Najahaty, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20413586&lokasi=lokal>

Abstrak

Pendeteksian topik merupakan suatu proses untuk mendapatkan topik dari suatu koleksi dokumen. Salah satu metode yang dapat digunakan untuk mendeteksi topik adalah nonnegative matrix factorization (NMF). Pada penelitian ini topik berita pada Twitter dideteksi menggunakan algoritma NMF berbasis metode langsung. Terdapat tiga tahap untuk menyelesaikan NMF berbasis metode langsung, yakni membentuk matriks kookurensi kata-kata, mencari kata anchor dan mencari matriks kata-topik. Pendeteksian topik dilakukan dengan jumlah topik yang berbeda-beda. Setelah didapatkan hasil berupa topik-topik dari pembicaraan di Twitter, kemudian tingkat akurasi topik-topik tersebut dianalisis menggunakan satuan topic recall, term precision dan term recall. Jumlah topik yang berbeda akan mempengaruhi tingkat akurasi topik-topik hasil NMF berbasis metode langsung. Pendeteksian topik merupakan suatu proses untuk mendapatkan topik dari suatu koleksi dokumen. Salah satu metode yang dapat digunakan untuk mendeteksi topik adalah nonnegative matrix factorization (NMF). Pada penelitian ini topik berita pada Twitter dideteksi menggunakan algoritma NMF berbasis metode langsung. Terdapat tiga tahap untuk menyelesaikan NMF berbasis metode langsung, yakni membentuk matriks kookurensi kata-kata, mencari kata anchor dan mencari matriks kata-topik. Pendeteksian topik dilakukan dengan jumlah topik yang berbeda-beda. Setelah didapatkan hasil berupa topik-topik dari pembicaraan di Twitter, kemudian tingkat akurasi topik-topik tersebut dianalisis menggunakan satuan topic recall, term precision dan term recall. Jumlah topik yang berbeda akan mempengaruhi tingkat akurasi topik-topik hasil NMF berbasis metode langsung.

<hr>

Topic detection is a process to get the topic of a collection of documents. One method that can be used to detect the topic is nonnegative matrix factorization (NMF). In this research, the topic of the news on Twitter detected using NMF algorithm based on the direct method. There are three stages to complete NMF-based direct method, they are, form the word-word coocurrence matrix, look for the anchor word and seek word-topic matrix. Topic detection performed by the different numbers of topic. Once the results are obtained in the form of topics of conversation in Twitter, then the level of accuracy of these topics were analyzed using the unit topic recall, term precision and term recall. Number of different topics will affect the accuracy of topics results of NMF-based direct method.