

Penentuan kata anchor berbasis singular value decomposition pada metode separable-nonnegative matrix factorization untuk pendeteksian topik = Finding anchor words based on singular value decomposition on separable nonnegative matrix factorization method for topic detection / Ika Dwi Novitasari

Ika Dwi Novitasari, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20455982&lokasi=lokal>

---

Abstrak

**ABSTRAK**

Pendeteksian topik merupakan suatu proses untuk menemukan topik atau pokok pembahasan utama dalam suatu kumpulan dokumen. Pada penelitian ini, pendeteksian topik diterapkan pada media sosial, yaitu Twitter. Pendeteksian topik pada Twitter secara manual sulit dilakukan karena terlalu banyak tweets. Oleh karena itu, dibutuhkan adanya pendeteksian topik secara otomatis. Salah satu metode otomatis untuk pendeteksian topik adalah metode Separable-Nonnegative Matrix Factorization S-NMF dengan algoritma AGM. S-NMF merupakan model berbasis faktorisasi matriks yang dapat diselesaikan secara langsung dengan menggunakan asumsi bahwa setiap topik memiliki satu kata yang tidak terdapat pada topik lainnya yang disebut kata anchor. S-NMF dengan algoritma AGM terdiri dari tiga tahapan, yaitu pembentukan matriks coocurance, penentuan kata anchor, dan recover. Dalam penelitian ini dilakukan proses penentuan kata anchor berbasis Singular Value Decomposition SVD . Kemudian, hasilnya akan dibandingkan dengan penentuan kata anchor metode Convex Hull berbasis Gram-Schmidt. Penelitian memberikan hasil bahwa dengan memperhatikan semua kata sebagai kandidat kata anchor, SVD memberikan hasil yang lebih baik daripada Convex Hull. Sedangkan, jika menggunakan anchor threshold, Convex Hull masih memberikan hasil yang lebih baik dibandingkan dengan SVD.

---

**ABSTRACT**

Topic detection is a process to find main topic or main subject of discussion in a collection of documents. In this research, topic detection is applied to social media, namely Twitter. Manual detection of topics on Twitter is difficult because of too many tweets. Therefore, it is necessary to detect topics automatically. One of the automatic methods for topic detection is the Separable Nonnegative Matrix Factorization S NMF method with the AGM algorithm. S NMF is a matrix factorization based model that can be solved directly using the assumption that each topic has one word that is not present in another topic called anchor words. S NMF with AGM algorithm consists of three stages, namely the formation of coocurance matrix, finding the anchor words, and recover. In this research, the process of finding anchor words was done based on Singular Value Decomposition SVD . Then, the result was compared to anchor word finding by Convex Hull based method. The results has shown that by considering all words as anchor word candidates, SVD gave better results rather than Convex Hull. Meanwhile, when the anchor finding was done by using anchor threshold, Convex Hull still gave better result rather than SVD.