

Penentuan kata-kata anchor pada metode separable nonnegative matrix factorization dengan menggunakan singular value decomposition dalam pendeteksian topik = Determination of anchor words on separable nonnegative matrix factorization method using singular value decomposition for topic detection

Delano Novrilianto, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20458891&lokasi=lokal>

Abstrak

ABSTRAK

Pendeteksian topik merupakan proses untuk mendapatkan topik dari koleksi data tekstual. Salah satu metode otomatis untuk masalah pendeteksian topik adalah Separable Nonnegative Matrix Factorization SNMF . Terdapat tiga tahap yang dilakukan untuk menyelesaikan SNMF yakni membentuk matriks kookurensi kata-kata, menentukan kata-kata anchor, dan mencari matriks kata-topik. Metode yang umum digunakan untuk menentukan kata-kata anchor pada tahap kedua dari penyelesaian SNMF adalah dengan metode berbasis Convex Hull. Pada penelitian ini digunakan pendekatan lain untuk menentukan kata-kata anchor yaitu dengan memakai metode Singular Value Decomposition SVD . Topik-topik yang dihasilkan dengan kata-kata anchor -nya ditentukan dengan metode SVD dievaluasi tingkat intepratabilitasnya dengan memakai satuan Pointwise Mutual Information PMI dan dibandingkan dengan topik-topik dimana kata-kata anchor -nya ditentukan dengan metode berbasis Convex Hull. Waktu komputasi yang dibutuhkan untuk menentukan kata-kata anchor dengan metode SVD juga dibandingkan dengan waktu komputasi yang dibutuhkan untuk menentukan kata-kata anchor dengan metode berbasis Convex Hull.

<hr>

ABSTRACT

Topic detection is the process of getting topics from a collection of textual data. One of the methods for detection problems is the Separable Nonnegative Matrix Factorization SNMF . There are three stages done to complete SNMF that is to form the word kookurensi matrix, determine the anchor words, and search for the word topic matrix. The common method used to determine the anchor words in the second stage of SNMF completion is the Convex Hull based method. In this research another approach is used to determine the anchor words, that is using Singular Value Decomposition SVD method. The resulting topics where the anchor words are determined by the SVD method will be evaluated for their interpretability level by using the Pointwise Mutual Information PMI unit and will compare with the topics where the anchor 39 s words are based on the Convex Hull based method. The computational time required to determine the anchor words by the SVD method will also be compared with the computational time required to determine the anchor words by the Convex Hull based method.