

Kernelisasi fuzzy c - means berbasis ruang eigen untuk pendeteksian topik pada berita online berbahasa Indonesia = Kernelization of eigenspaced ?? based fuzzy c - means for topic detection on Indonesian online news

Muktiari, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20477240&lokasi=lokal>

Abstrak

ABSTRAK

Pendeteksian topik adalah metode praktis untuk menemukan topik pada suatu koleksi dokumen. Salah satu metodenya adalah metode berbasis clustering yang mana centroid merepresentasikan topik contohnya eigenspace ndash; based fuzzy c ndash; EFCM . Proses clustering pada metode EFCM diimplementasikan pada dimensi yang lebih kecil yaitu ruang eigen. Sehingga akurasi dari proses clustering memungkinkan berkurang. Pada tesis ini, penulis menggunakan metode kernel sehingga proses clustering tersebut dapat diimplentasikan pada dimensi yang lebih tinggi tanpa mentransformasikan data ke ruang tersebut. Simulasi penulis menunjukkan bahwa kernelisasi ini meningkatkan akurasi dari EFCM berdasarkan skor interpretability pada berita online berbahasa Indonesia.

ABSTRACT

Topic detection is practical methods to find a topic in a collection of documents. One of the methods is a clustering based method whose centroids are interpreted as topics, i.e., eigenspace based fuzzy c means EFCM . The clustering process of the EFCM method is performed in a smaller dimensional Eigenspace. Thus, the accuracy of the clustering process may be reduced. In this thesis, we use the kernel method so that the clustering process is performed in a higher dimensional space without transforming data into that space. Author simulations show that this kernelization improves the accuracies of EFCM in term of interpretability scores for Indonesian online news.