

Abstract

Annotations on text documents such as articles, text scientific papers or other written materials may help the reader to more easily understand the contents of the text document. But giving the current annotation is done by the authors of the document text by choosing a category that is part of the recommendation annotations manually. This way is certainly not practical and require high accuracy of the author. Therefore in this thesis, an analysis and design process in the form of annotations on the category automatically. Method which is the main focus of this research is the Latent Semantic Analysis (LSA). Where will be analyzed also optimize the use of LSA in its application to the process of recommendation based annotation of text documents. LSA is a technique of mathematical / statistical relationships to extract and infer contextual meaning of the word is applied to the required text. This method has the characteristic to extract and represent sentences with mathematical calculations and the importance of key words contained in a sentence. Mathematical calculations done by mapping the presence or absence of the group said on semantic matrix and then processed using linear algebra techniques Singular Value Decomposition (SVD). Application of LSA method will be equipped with a text preprocessing consisting of tokenizing, filtering and stemming and words weighting to the application of term frequency.invers document frequency (tf.idf) weighting methods, which all of these methods will be used also in the process recommendation annotation.

Keywords : *Annotation recommendation, Latent Semantic Analysis, term frequency.invers document frequency, Text Processing.*