Text Summarization and Singular Value Decomposition

Josef Steinberger[1], Karel Ježek[1]

[1] University of West Bohemia in Pilsen,
Department of Computer Science and Engineering,
30614, Univerzitni 22, Plzeň, Czech Republic
{jstein, jezek_ka}@kiv.zcu.cz

Abstract. In this paper we present the usage of singular value decomposition (SVD) in text summarization. Firstly, we mention the taxonomy of generic text summarization methods. Then we describe principles of the SVD and its possibilities to identify semantically important parts of a text. We propose a modification of the SVD-based summarization, which improves the quality of generated extracts. In the second part we propose two new evaluation methods based on SVD, which measure content similarity between an original document and its summary. In evaluation part, our summarization approach is compared with 5 other available summarizers. For evaluation of a summary quality we used, apart from a classical content-based evaluator, both newly developed SVD-based evaluators. Finally, we study the influence of the summary length on its quality from the angle of the three evaluation methods mentioned.