

WEBSOM Method - Word Categories in Czech Written Documents



MOUŠEK, R., MAUTNER, P. WEBSOM Method - Word Categories in Czech Written Documents. In *Text, Speech and Dialogue*. Berlin: Springer, 2009. s. 85-92. ISBN: 978-3-642-04207-2

We applied well-known WEBSOM method (based on two layer architecture) to categorization of Czech written documents. Our research was focused on the syntactic and semantic relationship within word categories of word category map (WCM). The document classification system was tested on a subset of 100 documents (manual work was necessary) from the corpus of Czech News Agency documents. The result confirmed that WEBSOM method could be hardly evaluated because humans have problems with natural language semantics and determination of semantic domains from word categories.

13.09.2009

Plzeň