

The compression of texts written in natural language can exploit information about its linguistic character. It is shown that separation of coding of part-of-speech tags of a sentence (type of sentence) from the text alone can improve resulting compression ratio. For this purpose the tagging method NNTagger based on neural networks is designed. This thesis is focused on specification of the compression model of texts written in Czech. We propose methods of constructing of initial dictionaries and test their influence on the compression ratio.