

Při kompresi textu můžeme s úspěchem využít znalosti o přirozeném jazyce. V předkládané práci poukazujeme na možnost kódovat samostatně posloupnost slovních druhů nacházejících se ve větě (větný typ) a text samotný. Za tímto účelem je navrhnutá metoda značkování založená na neuronových sítích (NNTagger). Dále se práce soustřeďuje na specifikaci takového modelu komprese českých textů.

Navrhujeme zde metody vytvoření jejich inicializačních slovníků a testujeme vliv na výslednou kompresi.