

This thesis focuses on the task of automatic morphological inflection of Czech nouns, specifically in out-of-vocabulary (OOV) conditions (inflecting previously unseen words). We automatically extracted a large dataset suitable for training and evaluation in the OOV conditions. We also manually built a real-world OOV dataset of neologisms. We developed three different systems: a retrograde model performing a variation of kNN algorithm, and two sequence-to-sequence (seq2seq) models based on LSTM and Transformer. Compared to an available rule-based inflection system sklonuj.cz and standard SIGMORPHON shared task baselines, our seq2seq model reaches the best results in the standard OOV conditions. Moreover, it achieves state-of-the-art results for 6 out of 16 development languages from SIGMORPHON 2022 shared task data in the OOV evaluation (feature overlap) on large data condition. On the real-world OOV dataset, the retrograde model outperforms all neural models and is competitive with a non-neural SIGMORPHON baseline. We release the inflection system with seq2seq model as a ready-to-use Python library. It could serve as a complement to the state-of-the-art dictionary-based inflection system MorphoDiTa as a back-off for OOV words, especially once extended to other parts of speech.