

Tato práce porovnává dva přístupy k automatické předanotaci sémantických tříd sloves ve větách za účelem přidání nového jazyka do ontologie SynSemClass. Oba přístupy vycházejí z vícejazyčného deep learning klasifikačního modelu, který byl fine-tunovaný na již anotovaných anglických, českých a německých datech z ontologie. První, více tradiční, přístup je annotation projection. Používá paralelní korpus a výše zmíněný model k vytvoření predikcí na zdrojovém jazyce, který je již obsažen v ontologii, a tyto predikce projektuje na cílový jazyk pomocí automatického word alignmentu. Druhý přístup, zero-shot cross-lingual transfer, předpokládá, že vícejazykové schopnosti deep learning modelu jsou dostatečné a že můžeme vytvořit kvalitní predikce přímo na cílovém jazyce, i když model nebyl nikdy trénován pro danou úlohu na daném cílovém jazyce. Pro účely vyhodnocení ručně vytváříme a anotujeme malý korejský dataset za účelem otestování výsledků na jazyce, který se významně liší od angličtiny, češtiny a němčiny. Dospíváme k závěru, že zero-shot transfer vykazuje výrazně lepší výkon než annotation projection ( $p < 0,005$ ), s hodnotami recall a precision 0,54, ve srovnání s 0,37 recall a 0,41 precision u annotation projection. Také provádíme analýzu chyb a zjišťujeme, že dodatečné kroky annotation projection zavádějí kaskádovité chyby a že volný překlad sám o sobě představuje problém.