

Předtrénované jazykové modely jako například BERT v posledních letech sklízí mnoho úspěchů, je však stále obtížné uplatnit je v úlohách generování textů přirozeného jazyka. Tato práce se zabývá nedávno navrženou technikou tzv. adaptérů jako slibné alternativy k dotrénování celé předtrénované sítě a studuje ji v oblasti strojového překladu. Adaptéry umožňují dotrénovat jen velmi malou část předtrénované sítě. Ukazujeme, že s vhodnou inicializací dosahují adaptéry lepších výsledků než trénování modelů od počátku; s adaptéry se přitom trénuje podstatně méně vah, než má plný model. Překvapující zjištění je, že adaptéry dovolí dosáhnout kvality blízké základnímu modelu i v případě, že je vložíme a natrénujeme do sítě s fixními váhami, které byly nastaveny náhodně. Předtrénování základní sítě na velkých datech si v tomto případě můžeme ušetřit. Dále zkoumáme účinnost adaptérů v architektuře Transformeru uplatněné v úloze strojového překladu. Adaptéry vkládáme jen do enkodéru nebo jen do dekodéru, a pro ušetření paměti GPU se rovněž pokoušíme zmenšit velikost předtrénovaného modelu. Ukazuje se, že adaptéry použité jen v enkodéru vedou k obdobné kvalitě jako adaptéry použité v enkodéru i dekodéru současně. Studie zmenšování modelu pak ukazuje, že použití pouze poloviny předtrénovaných vah může s adaptéry vést ke zlepšení výsledné kvality.