

Posudek na teze disertace Dušana Variše „Exploring Human-like Learning Capabilities of Neural Machine Translation“

Autor posudku: Ondřej Dušek

14. 2. 2020

Shrnutí obsahu: Předmětem disertace Dušana Variše je zkoumání schopností neuronových sítí na úloze strojového překladu v oblastech, kde schopnosti současných neuronových architektur strojového učení výrazně zaostávají za lidskými schopnostmi učit se. Autor jmenuje tyto čtyři oblasti:

1. kontinuální učení (schopnost naučit se nové dovednosti a nezapomenout předchozí),
2. skládání znalostí (schopnost rozložit složitou úlohu na několik jednodušších, vyřešit je a složit výsledky),
3. učení s omezenými zdroji (schopnost učit se z mála trénovacích příkladů),
4. učení se multimodální reprezentace znalostí (kombinace znalostí z textu a obrazu).

Po úvodním představení problému teze v následující sekci shrnují relevantní literaturu ke všem čtyřem problémům. V další sekci se autor věnuje formální definici neuronových architektur pro jazykové modelování a strojový překlad; k tomu přidává krátký popis evaluačních metrik. Čtvrtá sekce obsahuje popis experimentů, které autor dosud během svého výzkumu provedl. Ty jsou sdruženy podle jednotlivých publikací a doplněny o dosud nepublikované výsledky; autor vždy krátce popíše problém a zvolený přístup a shrne hlavní poznatky. Výsledky zahrnují práci v oblastech automatické posteditace strojového překladu, multimodálního strojového překladu, předtrénování pro strojový překlad s omezenými zdroji a učení multimodální reprezentace znalostí. Předposlední sekce tezí představuje autorovy další výzkumné plány, týkající se zejména vývoje nové modulární neuronové architektury založené na vícehlavém mechanismu attention, která by měla být schopna přepínat mezi různými sadami naučených parametrů v závislosti na vstupu. Autor navrhuje dvě varianty architektury a přidává seznam experimentů, kterým ji chce podrobit. Nechybí ani „plán B“ – možné další architektury, které by autor mohl prozkoumat, pokud plánové experimenty skončí neúspěchem. Teze končí krátkým shrnutím.

Přednosti:

- Problémy navržené ke zkoumání považuji za důležité a vhodně zkombinované.
- Strojový překlad a jeho varianty (multimodální/vícezdrojový/s omezenými zdroji) jsou jednoznačně vhodná testovací úloha pro zkoumání těchto problémů.
- Navrhované experimenty vypadají slibně; myslím, že představená modulární architektura by mohla vést ke zlepšení alespoň v některých ze zkoumaných oblastí.
- Je dobré, že autor přemýšlí o alternativních možnostech experimentů pro případ, že navrhovaná architektura nepřinese pozitivní výsledky.
- Shrnutí relevantní literatury k tématu je přehledné, rozsahem adekvátní, dobře čitelné a až na nepodstatné detaily k němu nemám výhrady.

Nedostatky:

- Teze nemají explicitně formulované cíle práce, hypotézy a/nebo výzkumné otázky. Ty se samozřejmě můžou v průběhu další práce změnit, ale přítomnost alespoň nějakých pracovních verzí by byla vhodná.
- Výsledky experimentů nejsou moc dobře vztaženy k oblastem zájmu disertace. Kdybych četl jen sekci s výsledky, došel bych asi k závěru, že předmětem disertace je jakkoliv vylepšit strojový překlad zkoušením různých neuronových architektur.
- Navrhované evaluační metriky nepovažuji za dostatečné pro zhodnocení přínosů nových architektur. Sice se s autorem shodnu na tom, že manuální evaluace je drahá, těžce reprodukovatelná a nevhodná pro vývoj, ale nedokážu si představit její absenci při finálním vyhodnocení.

Automatické metriky pro strojový překlad mají mnoho známých problémů (viz např. Novikova et al., 2017; Reiter, 2018) a snadno mohou podat zkreslenou představu o schopnostech jednotlivých systémů.

- Ač je práce psaná většinou dobře srozumitelným jazykem, organizace textu je místy trochu chaotická:
 - Čtvrtá z navrhovaných oblastí je nelogicky oddělená od zbývajících tří jako nerovnocenná, přitom v dalším textu se její pojednání od ostatních tří moc neliší.
 - Oblasti zájmu jsou většinou definovány jako schopnosti (kompozice znalostí, učení s omezenými zdroji), z toho se vymyká „katastrofické zapomínání“ definované jako problém, který ale autor v dalším textu vhodněji opisuje jako schopnost „kontinuálního učení“.
 - Třetí sekce s metodologií není příliš propojená se zbytkem textu a obsahuje hlavně základní definice, takže v tezích je spíše zbytečná (v samotné disertaci ale samozřejmě bude žádoucí). Její uspořádání taky není ideální, zasloužila by víc podsekcí.
 - Dělení experimentů na publikované a nepublikované je pro účel tezí zbytečné.
- Některé detaily – zejména struktura architektur a dalších přístupů zmiňovaných v sekcích o dosavadních i plánovaných experimentech – by zasloužily podrobnější nebo přiléhavější vysvětlení. Jako čtenář, který se orientuje v NLP i hlubokém učení, ale nezná všechny nové trendy ve strojovém překladu a jazykovém modelování, jsem nucen některé věci buď nepochopit, nebo dohledávat (např. metriky pro evaluaci kontinuálního učení, trénovací techniky od Lamplea et al., elastickou konsolidaci vah, chybové funkce Bordese et al. a další).

Celkové hodnocení: Jak vyplývá z předchozího textu, přednosti se týkají zejména tématu a návržení samotné práce, nedostatky pak hlavně jeho zpracování v tezích. Rozhodně si myslím, že na základě těchto tezí a plánovaných experimentů může vzniknout úspěšná disertace. Na druhou stranu by ale explicitní (aspoň pracovní) seznam cílů, hypotéz a výzkumných otázek byl namísto a celkem rád bych ho viděl už během obhajoby tezí. Dále by bylo dobré, aby autor lépe vztáhl své současné experimenty k deklarovaným oblastem zkoumání a pokusil se odpovědět na otázky typu: „Co znamenají výsledky mých experimentů s posteditací pro schopnost kontinuálního učení?“ Rád bych taky věděl, jaké má autor časové plány pro další experimenty.

Reference:

Novikova, Jekaterina, et al. “Why We Need New Evaluation Metrics for NLG.” *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, 2017, pp. 2243–53. <http://aclweb.org/anthology/D17-1238>

Reiter, Ehud. “A Structured Review of the Validity of BLEU.” *Computational Linguistics*, vol. 44, no. 3, Sept. 2018, pp. 393–401. <https://www.aclweb.org/anthology/J18-3002/>