

## Review of Master's Thesis

**Student:** Slávka Michal, Bc.  
**Title:** Multilingual Open-Domain Question Answering (id 23369)  
**Reviewer:** Dočekal Martin, Ing., DCGM FIT BUT

- 1. Assignment complexity** **more demanding assignment**  
Pro řešení bylo nutné nastudovat větší množství odborných článků z poslední doby a dobře se orientovat v problematice umělých neuronových sítích.
- 2. Completeness of assignment requirements** **assignment fulfilled**  
Zadání této práce bylo splněno.
- 3. Length of technical report** **within minimum requirements**  
Textová zpráva splňuje minimální počet normostran. Některé sekce jsou příliš stručné. Především popis metriky BLEU.
- 4. Presentation level of technical report** **60 p. (D)**  
Jednotlivé kapitoly jako celek na sebe dobře logicky navazují. Nicméně některé pasáže mohou být pro čtenáře méně srozumitelné či zavádějící. Například již zmiňovaná sekce o použitých metrikách je příliš stručná. V sekci 3.2 je zmíněn seznam použitých heuristik, který se tváří jako úplný, ale není. Dále v sekci 4.2.1 autor tvrdí, že daná loss je nenáročná na výpočet, ale není zřejmé proč.  
  
Bohužel se v práci objevují i faktické nesrovnalosti. V úvodu práce autor tvrdí, že seq2seq modely se používají v extraktivním přístupu k úloze odpovídání na otázky, ale ty se běžně používají pro generativní. V sekci 2.1 je tvrzení, že dataset MKQA obsahuje člověkem přeložené otázky i odpovědi, přestože původní autoři datasetu zmiňují, že pro odpovědi je použito i automatického překladu. U popisu datasetů je uváděn jiný počet zahrnutých jazyků. Rovnice 4.11 má chybnou levou stranu. Obrázek 6.1, obsahující graf, nemá popsanou jednotku u ypsilonové osy. Dále se zde objevují i některé další problémy, které jsou příbuzné zde vyjmenovaným.
- 5. Formal aspects of technical report** **70 p. (C)**  
Práce obsahuje větší množství zdvojených slov (was was) a objevilo se i zdvojení myšlenky (positionally invariant order-independent). Nadpisy 2.3 a 4.3.2 obsahují mezery navíc. Obrázek 6.6 není nikde odkazován z textu. Odkaz na straně 4 v 2. poznámce pod čarou je nefunkční vlivem přebytečné tečky. Celkově však formátování textových pasáží je na dobré úrovni, až na rozšířený abstrakt, kde dochází k přetékání textu. Autor vhodně doplňuje text o ilustrace a rovnice.
- 6. Literature usage** **60 p. (D)**  
Autor vybral množství relevantních zdrojů, které se z velké části skládají z odborných článků. Většinou jsou tvrzení v práci jasně podložena patřičným zdrojem, ovšem (spíše výjimečně) v některých částech citace chybí. Bohužel došlo v práci k porušení citační etiky, když autor neuvedl zdroj u převzatých obrázků 3.1 a 5.1. Dále tabulka 4.1 uvádí výsledky, které se liší od těch uvedených v citovaném zdroji a namísto toho obsahuje výsledky ze starší verze preprintu.
- 7. Implementation results** **90 p. (A)**  
Za realizační výstup považuji jak implementovaný systém, tak provedené experimenty. Autor vytvořil systém, který je použitelný pro problém vícejazyčného odpovídání na otázky a umožňuje experimentovat s popsánými modely. Práce obsahuje velké množství zajímavých experimentů, které jsou i náležitě okomentovány.
- 8. Utilizability of results**  
Provedené experimenty nastiňují problémy a výhody, které přináší podpora vícejazyčnosti pro automatické odpovídání na otázky, lze tedy na tyto výsledky navázat v další práci.
- 9. Questions for defence**
  - Jak si vysvětlujete, že se lepší výkon překladače, hodnoceného podle BLEU, neprojevuje na lepších výsledcích, přestože přímo použití lidského překladu pozitivní vliv má.
- 10. Total assessment** **70 p. good (C)**  
Autor provedl velké množství zajímavých experimentů, ale bohužel text práce působí celkovým dojmem, že byl napsán narychlo.

In Brno 27 May 2021

Dočekal Martin, Ing.  
reviewer