

Posudek rigirózní práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce Kateřina Macková
Název práce Question Answering in Czech via Machine Translation and Cross-lingual Transfer
Rok odevzdání 2022
Autor posudku David Mareček **Role** oponent
Pracoviště ÚFAL MFF UK

Text posudku:

Rigorózní práce prozkoumává možnosti aplikace úlohy odpovídání na otázky (question answering) na češtinu. Hlavním problémem je zde neexistence českých dat. Práce se tedy zabývá procesem automatického překladu anglického datasetu do češtiny. Dále pak porovnává několik možností trénování modelů a to 1) trénování modelu na originálních anglických datech a jeho přímá aplikace na český vstup, 2) použití anglického modelu a automatický překlad vstupů/výstupů z/do češtiny, 3) natrénování modelu na češtině.

Byly provedeny originální experimenty ukazující sílu multilinguálních modelů v transfer-learningu, jejichž výsledky jsou srovnatelné s trénováním modelu na českých datech. Z experimentálního hlediska práce naplňuje požadavky diplomové a tedy i rigirózní práce.

Z hlediska kvality textu je však práce slabší. Je psána často méně srozumitelně, během čtení zůstává mnoho věcí nejasných nebo nevysvětlených. Některé kusy textu se v práci zbytečně opakují. Některé části mi přijdou napsané dokonce velmi zavádějícím až nepravdivým způsobem, ze kterého mi připadá, že autorka nechápe, jakým způsobem funguje neuronová síť, kterou používá (viz následující bod 3). Zde uvádím konkrétní nedostatky:

1. Třetí odstavec v kapitole 1 (strana 4) je téměř totožný se čtvrtým odstavcem sekce 2.3.4 (strana 11). Statistiky o výsledném datasetu jsou shodně na stránkách 5, 35, a 40.
2. Sekce 2.3 je napsaná hodně vágně a není jasné, zda popisuje obecně všechny modely, nebo jen ty neuronové, nebo jen Transformery, výsledný mix informací však nepasuje ani na jednu ze mnou uvedených možností.
3. Kapitola 4 popisuje modely používané v úloze “Question Answering”. Zejména ale z popisu architektury BERT v sekci 4.2 mám jisté pochybnosti, zda autorka práce rozumí tomu, jak síť funguje. Její popis “attention” mechanismu na straně 26, který je popsán jako dvousměrný, kde jeden směr spojuje slova dotazu se slovy textu a druhý směr naopak, mi připadá jako velmi zavádějící. U tohoto popisu autorka navíc odkazuje na obrázek 4.4

(převzatý z původního článku od Vaswaniho), který se sice týká attention mechanismu, ale tyto dva směry vůbec nijak neilustruje. Attention mechanismus v této síti jednoduše propojuje mezi vrstvami všechny tokeny mezi sebou. Je samozřejmě možné, že výsledná síť se naučí některé vzory, ale to není vlastnost architektury. Další zásadní nepřesnost je v sekci 4.2.2, kde autorka tvrdí, že tokeny odpovídají slovům a vůbec nezmiňuje fakt, že transformer pracuje na úrovni podslov (subwords). Dále, u popisu trénování transfromeru, které je rozděleno na dvě části (úlohy) autorka odkazuje na obrázek 4.5, který však nic takového neilustruje.

4. Obrázky 5.1 a 5.2 ilustrující formát datasetu mi připadají špatně. Myslím, že tagy “question” a “id” by měly být na stejné úrovni jako “answers”.
5. Nepochopil jsem vůbec příklady v sekci 5.1.3. Není jasné, zdali se týkají výsledků před použitím přepočítání indexu, nebo až poté.

Naopak, kapitolu 6 hodnotím velmi kladně, sumarizace a diskuse výsledků ukazuje jasný výsledek experimentů, který podtrhuje sílu transfer-learningu použitím multilinguálních modelů. Zde mám pouze jednu poznámku a to, že není jasné, zda anglické modely evaluované na češtině také používaly pro vyhodnocení lematizaci a Derinet.

Závěr

Rigorózní práce Kateřiny Mackové je po experimentální stránce velmi dobrá a obsahuje jasné a dobře diskutované závěry. Kvalita textu práce je však velmi slabá, a ukazuje například, že studentka modelu, který v práci používá, příliš nerozumí. Jako diplomová práce by podle mého názoru tato práce asi měla být obhájena a nechávám na komisi, zda by měla projít i jako práce rigorózní.

V Praze dne 30. 9. 2022

Podpis: