

Posudek oponenta diplomové práce

Jméno a příjmení autora posudku: Vladislav Kuboň

Jméno a příjmení autora práce: Matouš Macháček

Název práce: Measures of Machine Translation Quality

Vlastní text (sem prosím napište text posudku, délka textu posudku není omezena):

Diplomová práce Matouše Macháčka se zabývá problémem měření kvality automatického překladu, je napsána anglicky. Skládá se ze šesti kapitol a je doplněna datovým nosičem.

Diplomová práce má v podstatě dvě hlavní části. V té první autor popisuje svou vlastní novou metodu měření kvality automatického překladu, druhá část (kapitola 4) potom poskytuje přehled jiných metod, použitých v soutěži o nejlepší metriku v rámci pravidelného workshopu automatického překladu WMT14. Z hlediska struktury práce by asi bylo přirozenější nejprve seznámit čtenáře s existujícími metrikami a teprve poté navrhnout novou, takto není při čtení práce zpočátku jasné, jaké existující metriky autorovi posloužily jako inspirace.

Podstatou autorem navrhované metriky SegRanks je přesvědčení, že rozdělení přeložených vět na menší lingvisticky motivované segmenty pomůže zjemnit hodnocení kvality. Hodnotitelé dostanou místo dlouhých vět kratší úseky, které bude jednodušší seřadit podle kvality překladu. Autor vychází ze své zkušenosti s hodnocením kvality překladu ze dvou ročníků WMT, kde se často stávalo, že u dlouhých vět byla jedna část lepší a jiné horší a bylo tedy velmi obtížné celé věty seřadit.

Autor si je také vědom možných nedostatků navrhované metody, uvádí celkem čtyři důvody, kvůli nimž by mohla selhávat. Všechny jsou závažné, a jak se ukáže v dalších kapitolách práce, mají svůj podíl na tom, že navrhovaná metoda nespĺňuje očekávání. Je proto škoda, že své rozhodnutí zvolit právě tuto metodu autor nepodpořil dalšími argumenty.

Kapitola 2 dále popisuje anotační nástroj, pomocí kterého jednotliví hodnotitelé řadí přeložené segmenty. Nástroj je uživatelsky přívětivý, má jednoduché ovládání a umožňuje pohodlné hodnocení. Ve srovnání se standardně používaným prostředkem hodnocení Amazon Mechanical Turk se po uživatelské stránce jedná o značné zlepšení uživatelského komfortu. Reprezentace získaných dat je efektivní a dostatečně jednoduchá. Tento nástroj patří mezi důležité výsledky práce a měl by být dán k dispozici odborné veřejnosti např. prostřednictvím stránek projektu LINDAT.

Kapitola 3 obsahuje důkladný popis experimentů s novou metrikou. Autor rozebírá jak její možné využití při vývoji a ladění systémů automatického překladu, tak i pro hodnocení kvality nových přeložených textů. V této části práce autor prokazuje, že se v dané oblasti velmi dobře vyzná, že umí sestavit a metodicky správně vyhodnotit experimenty. Kapitola obsahuje značné množství tabulek, které jsou v textu podrobně rozebrány a vysvětleny. Výsledky ukazují, že metoda SegRanks vede k lepší mezianotátorské shodě, dokáže pomoci při ladění systémů automatického překladu, ovšem zároveň také v podstatě potvrzují, že není vhodná pro hodnocení nových přeložených textů.

Druhé téma práce, přehled existujících metrik a rozbor jejich výsledků, je koncentrováno do kapitoly 4. Určitým nedostatkem této kapitoly je to, že se nezabývá obecně existujícími metrikami, bere v úvahu pouze ty, které se přihlásily do soutěže Metrics Shared Task v rámci WMT 14, kterou autor spolupořádal se svým vedoucím, dr.Bojarem. Velké části textu této kapitoly byly doslova přejaty z článku Macháček, Bojar: Results of the WMT 14 Shared Metrics Task, což v práci není výslovně uvedeno. Vzhledem k tomu, že původní článek má dva spoluautory, není jasné, do jaké míry je obsah této kapitoly vlastní prací autora. Autor by měl při obhajobě podíl své práce vyjasnit. Po obsahové stránce však není přehledu metrik co vytknout, autor jednotlivé metriky nejprve popisuje a potom prostřednictvím celé řady tabulek i porovnává.

V závěru pak autor ještě jednou shrnuje dosažené výsledky a navrhuje, jak by bylo možné metodu SegRanks dále zlepšovat. Tyto návrhy vycházejí z výsledků práce a jsou logické.

Po formální stránce je třeba vyzdvihnout bohatou bibliografii, pečlivé typografické zpracování a dobrou angličtinu prakticky bez chyb.

Z hlediska celkového hodnocení práce je nutné konstatovat, že autor splnil zadání, provedl značné množství experimentů a dokázal je správně vyhodnotit a interpretovat. Je škoda, že přestože v závěru navrhuje několik dalších možností zlepšení metody SegRanks, nedokázal autor alespoň některé z těchto zlepšení implementovat tak, aby metoda skutečně prokázala svou užitečnost i při hodnocení nových překladů.

Doporučení k obhajobě:

Z výše uvedených důvodů práci *doporučuji* k obhajobě.

Vynikající práce vhodná pro soutěž studentských prací	ANO <input type="checkbox"/>
---	------------------------------

Seznam soutěží studentských prací, viz <http://www.mff.cuni.cz/studium/bcmgr/prace/>

Pokud jste výše zaškrtnli ANO, zdůvodněte prosím svůj návrh, případně uveďte konkrétní soutěž, pro kterou je práce vhodná (rámeček lze nechat prázdný, pokud za dostatečné zdůvodnění považujete text posudku):

--

V Praze dne: 1.9.2014

Podpis:**