

## Posudek vedoucího diplomové práce

Jméno a příjmení autora posudku: Ondřej Bojar

Jméno a příjmení autora práce: Dušan Variš

Název práce: Automatic Error Correction of Machine Translation Output

---

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

Diplomová práce Dušana Variše svým cílem navazuje na úspěšný systém Depfix Rudolfa Rosy. Depfix pomocí souboru ručně sestavených pravidel opravoval výstupy automatického překladu z angličtiny do češtiny. Práce Dušana Variše má za cíl odstranit nebo alespoň výrazně omezit jak závislost na cílovém jazyce, tak i nutnost ručně formulovat pravidla.

Předkládaná práce je přehledně členěna do osmi kapitol v nichž postupně a beze zbytku řeší zadaný úkol. Po stručném seznámení s problematikou automatických posteditací shrnuje potřebné základní stavební kameny lingvistické analýzy (kap. 2) a dostupná trénovací data a jejich vlastnosti (kap. 3). V kapitole 4 následuje podrobná rozvaha a definice úlohy prostřednictvím dvou samostatných kroků: identifikace problematických míst a návrhy správných hodnot jednotlivých morfologických kategorií pro detekovaná problematická místa. V této kapitole je rovněž vyhodnocen horní odhad úspěšnosti při zvoleném způsobu využití trénovacích dat pro strojové učení. Zatímco dosažitelná přesnost se pohybuje mezi 79 % při použití ruční reference a 94 % při snaze replikovat Depfix, míra využití trénovacích vět (impact) je velmi malá, 6-8 %. Pro získání spolehlivých trénovacích dat se tak očekává velká náročnost na objem vstupních textů.

Návrhu a trénování samotných modelů je věnována kap. 5. Ve všech krocích o klasifikátory, přičemž autor experimentuje na třech datových sadách s několika typy klasifikátorů a zabývá se jak problémem nevyvážené trénovací množiny, kde převládají bezchybná slova, tak výběrem nejlepší množiny rysů vstupujících do klasifikátoru. Dílčí varianty metody jsou empiricky vyhodnoceny a výsledky řádně komentovány. Celkovým vyhodnocením systému se pak zabývají kapitoly 6 (čeština) a stručnější 7 (němčina). Nechybí ruční vyhodnocení na malém vzorku a výsledky naznačují, že systém je schopen díky svým opravám v češtině zlepšit asi 3/4 zasažených vět. Podobně jako u přípravy trénovacích dat je však celkově nízký počet zasažených vět (impact). Výsledky na německém prototypu jsou citelně horší (dosažitelná přesnost jen 34 %, využití trénovacích vět 3 %, závěrečná přesnost jen 13 %, i když podíl zasažených vět je vyšší, až 6 %). V češtině je zlepšení patrné i při využití automatického hodnocení překladu (BLEU), v němčině dochází „opravou“ naopak k poklesu kvality hypotéz. Obecně se ukazuje, že detekce problematických míst je výrazně náročnější úkol než predikce správné hodnoty morfologických kategorií, pokud je chybné slovo nalezeno. Jak autor sám komentuje, nízká kvalita německé varianty systému může být do určité míry dána zastaralým generátorem opravených německých slov.

Po formální stránce je práce kvalitní. Je psána anglicky s akceptovatelným množstvím gramatických chyb (např. neurčité členy v množném čísle), literatura je řádně citována, všechny tabulky a grafy řádně odkazovány a komentovány. Práce netrpí žádnými typografickými nešvary.

Svou diplomovou práci Dušan Variš prokázal schopnost samostatně řešit komplexní úlohy a přesně a srozumitelně rozvahu, postup i výsledky popsat.

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/>

V Praze dne: 26. srpna 2016

Podpis:\*\*

\* *nehodící se škrtněte (vymažte)*

\*\* *do SISu vkládejte formulář nepodepsaný (ve formátu PDF), podpis je potřeba doplnit až na vytištěný posudek.*