

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Generování přirozeného jazyka ze znalostních databází
Jméno autora:	Ondřej Kobza
Typ práce:	diplomová
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra počítačů
Oponent práce:	Ing. Jiří Spilka Ph.D.
Pracoviště oponenta práce:	CIIRC, ČVUT v Praze

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	náročnější
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Práce se zabývá generováním vět ze znalostních databází reprezentovaných pomocí Resource description framework (RDF). Datový rámec objekt – predikát – subjekt je použit pro tvorbu přirozeného jazyka s cílem využít takto generované věty k obohacení hlasového asistenta Alquist. Zadání práce hodnotím jako náročnější, protože vyžaduje orientaci v nejnovějších metodách zpracování přirozeného jazyka a práci s velkými modely.	

Splnění zadání	splněno
<i>Posuďte, zda předložená závěrečná práce splňuje zadání. V komentáři případně uveďte body zadání, které nebyly zcela splněny, nebo zda je práce oproti zadání rozšířena. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.</i>	
Zadání práce bylo splněno. Nad rámec zadání byly zvoleny dva možné přístupy, kde se student snažil najít kompromis mezi výslednou přesností modelu a rychlostí inference.	

Zvolený postup řešení	správný
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
Zvolený postup řešení považuji za správný. Student nejprve uvádí přehled dostupných datasetů a algoritmů, stručně je představuje a vybírá kandidáty pro experimenty. Dále také navrhuje vlastní heuristiku s cílem zlepšit přesnost modelů.	

Odborná úroveň	C - dobře
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
Práce má z mého pohledu příliš velký záběr a snaží se představit a popsat celé spektrum algoritmů. Na jednu stranu je tento výčet zajímavý, ale bohužel kvalita popisu jednotlivých algoritmů není dle mého názoru dostatečná. Zde bych jednoznačně preferoval kvalitu před kvantitou. Stejně tak je tomu s výběrem metriky pro hodnocení. Student se nedokázal rozhodnout, která metrika je nejdůležitější a tak v porovnání používá pět metrik. Což pochopitelně vede k tomu, že je nutné klikovat mezi výsledky, protože žádný model nepřevyšuje ostatní ve všech sledovaných metrikách. Dále k výsledkům, jedna skupina modelů je hodnocena na datasetu WebNLG 2017 a druhá na datasetu WebNLG 2020 a tak se jejich porovnání komplikuje. Interpretaci výsledků bych si také představoval detailnější.	

Formální a jazyková úroveň, rozsah práce	C - dobře
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
Práce je psaná v anglickém jazyce, ale bohužel není úplně čtivá. Text práce není plynulý a je příliš stručný. Zde se opět negativně projevuje velký rozsah práce. Některé zkratky nejsou vysvětleny (RDF, API, CTU, LSTM) a uvedení online zdrojů také není konzistentní. Některé jsou uvedeny pod čarou, některé v textu a další potom v použité literatuře.	

Výběr zdrojů, korektnost citací

B - velmi dobře

Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce. Charakterizujte výběr pramenů. Posuďte, zda student využil všechny relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.

Použitá literatura je adekvátní a odpovídající. K porušení citační etiky dle mého názoru nedošlo. Bohužel se v literatuře vyskytují zdroje, které neodpovídají citační normě, např. odkazy na online zdroje dbpedia.org, Bayu Distiawan et al. a dále celá řada dalších článků jako Dzmitry Bahdanau et al. Použitá literatura tak bohužel působí odbytým dojmem.

Další komentáře a hodnocení

Vyjádřete se k úrovni dosažených hlavních výsledků závěrečné práce, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, publikačním výstupům, experimentální zručnosti apod.

Vložte komentář (nepovinné hodnocení).

III. CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE

Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení. Uveďte případné otázky, které by měl student zodpovědět při obhajobě závěrečné práce před komisí.

Práce se zabývá zajímavým tématem a v principu student postupoval správně. Z experimentů a přiložených kódu je zřejmé, že student odvedl velké množství práce, ale kvalita práce utrpěla příliš ambiciózním rozsahem, který se negativně podepsal na kvalitě popisu a prezentaci výsledků.

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **C - dobře**.

Otázky do diskuze:

1. V úvodu se zmiňujete, že jedním z cílů je použít generování textu v konverzačním asistentu Alquist pro Alexa Prize 2021. Bylo Vaše řešení nakonec použito v systému Alquist?
2. V kontextu konverzačního asistenta a Alexa prize, jakou metriku byste zvolil k posouzení přínosu Vašeho řešení?
3. Mohl byste uvést několik reprezentativních příkladů generovaného textu. V práci to není uvedeno a je těžké si udělat představu, jaké kvalitě odpovídá průměrné skóre např. chr++ 0.661 pro T5-large v Tabulce 6.8.

Datum: **18.6.2021**

Podpis: