

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Tvorba dialogových aplikací na základě ontologie schema.org
Jméno autora:	Tatiana Okonechnikova
Typ práce:	diplomová
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra počítačů
Oponent práce:	Jan Pichl
Pracoviště oponenta práce:	Katedra kybernetiky

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	průměrně náročné
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Zadáním bylo prozkoumat technologie používané při tvorbě konverzační AI (chatbotů) a navrhnout poloautomatické generování dialogu s využitím strukturovaných znalostních databází a semistrukturovaných informací z webu. Aplikace měla být dále napojena na platformu Amazon Alexa nebo Google Home a otestována.	

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>	
Dle mého názoru se práce orientovala více na NLU části systému než na samotné generování dialogu ze znalostní databáze. Body zadání však byly splněny.	

Zvolený postup řešení	částečně vhodný
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
Menší výhrady mám ke zvolenému směru, kterým se práce vydává. Zadáním práce je (poloautomaticky) vytvořit konverzaci na základě externího zdroje znalostí. Předpokládal bych, že v práci bude podrobně zkoumáno a pospáno, jakým způsobem jsou data v těchto databázích uložena a jak je využít při tvorbě dialogu. Popisu těchto databází je v poměru k ostatním částem práce věnován malý prostor. Na druhou stranu je (v textu) věnován veliký důraz na NLU algoritmy (konkrétně rozpoznávání intentu a entit) na úkor zmiňovanému generování dialogu.	

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 se velmi podrobně zabývá teoretickým rozbohem NLP principů, kde začíná od samých základů popisovat ML principy od trénovací množiny, přes samotné trénování po cross-validaci. V implementační části je kladen důraz na detekci entit a rozpoznání intentu, ovšem kód pro vytvoření modelu pro intent není přiložen. Dále je popsáno získávání znalostí z databází a generování přirozeného jazyka pomocí bezkontextové gramatiky. V části experimentů je popsáno 5 proběhlých dialogů. K těmto konkrétním výsledkům by bylo dobré ještě doplnit hodnocení uživatelů a shrnutí (procentuální zastoupení) nalezených chyb.	

Formální a jazyková úroveň, rozsah práce	B - velmi dobře
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
Práce je na dobré jazykové úrovni, v textu se nevyskytují závažné chyby. Po typografické stránce lze vytknout umístění popisků pod tabulkou místo nad ní a chybné úvodní uvozovky u každého uvozaného textu. Teoretická část práce by nemusela být rozebírána od úplných základů, což by bylo vhodné s ohledem na délku práce (89 stran).	

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.

Studentka v práci využívá relevantní zdroje. Tvzení jsou řádně odcitována, výtku bych měl pouze k první větě v sekci 4.1, která je až na drobnosti stejná jako v blogu „An easy introduction to Natural Language Processing“ bez jakékoliv citace.

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 poloautomatickou tvorbou dialogu s využitím znalostních databází. V úvodní části se nachází popis chatbotů v souvislosti s chytrými asistenty. Následuje část týkající se znalostních databází a teoretická část popisující NLU a machine learning principy v NLP. Tato část popisující principy je psána učebnicovou formou a vysvětluje věci od samého základu (co je to trénovací a testovací množina), což je podle mého názoru zbytečné uvádět v diplomové práci v tomto oboru.

V implementační části je nejprve popsán systém jako celek a dále jsou pak rozebrány jednotlivé části. Z popisu rozpoznávání intentu (a dále pak v kapitole experimentů) vyplývá, že bylo experimentováno s několika architekturami neuronových sítí. V příloze k této práci se bohužel nenachází žádný kód určený k trénování těchto sítí, ale pouze jeho část, kde je hotový model načten ze souboru.

Celkově se bohužel nedá systém z příloženého zdroje otestovat. První nepříjemností je chybějící soubor závislostí. Toto se dá naštěstí překonat postupným procházením kódu a instalováním chybějících modulů. Nepřekonatelnou překážkou je však absence souborů obsahující natrénované modely.

Část práce popisující experimenty obsahuje 5 dialogů, kde za každým dialogem následuje analýza chyb. Jelikož podle zadání měly být dialogy testovány alespoň třemi různými uživateli, očekával bych nějaké uživatelské hodnocení subjektivního charakteru (zábavnost, srozumitelnost, ...).

Otázky:

1. Na straně 45 popisujete odstraňování stop-words a dokonce domain-specific stop-words. Používáte tuto metodu jen při extrakci informací pomocí prefixových stromů nebo i při použití StanfordNER?
2. Hledání klíčových slov s použitím Levenshteinovy vzdálenosti se hodí na eliminaci chyb v textových chatbotech. Jak byste řešila chyby v hlasových chatobech způsobené rozpoznáním řeči?
3. Jaké jsou manuální kroky pro implementaci jiné domény než filmy?
4. Z kódu na integraci s Amazon Alexa vyplývá, že všechny modely musí být znovu načteny před každou otáčkou dialogu. Jaké jsou časové prodlevy mezi jednotlivými dialogovými otáčkami na Alexe?

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

Datum: 28.1.2019

Podpis: