ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

POSUDEK VEDOUCÍHO ZÁVĚREČNÉ PRÁCE

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Porozumění nestrukturovanému textu a automatické odpovídání na dotazy
Jméno autora:	Bc. Jakub Konrád
Typ práce:	Diplomová
Fakulta/ústav:	Fakulta Elektrotechnická (FEL)
Katedra/ústav:	Katedra řídící techniky
Vedoucí práce:	Ing. Jan Šedivý CSc.
Pracoviště vedoucího práce:	ČVUT CIIRC

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání — Náročná

Hodnocení náročnosti zadání závěrečné práce.

Náročnost tématu je přiměřená typu závěrečné práce.

Splnění zadání — Splněno

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.

Zadání bylo beze zbytku splněno a demonstrováno na příkladech. Výsledky jsou uspořádány ve formě přehledné tabulky.

Zvolený postup řešení — vynikající

Posuďte, zda student zvolil správný postup nebo metody řešení.

Navržené řešení je detailně odůvodněno. Bez výhrad.

Odborná úroveň — A-výborně

Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.

Navržené řešení ukazuje jak použít transfer learning pro zvýšení přesnosti odpovídání dotazů z nestrukturovaného textu. Řešení navazuje na výzkum v oblasti odpovídání složitých dotazů a ukazuje jak postupovat při změně domény resp. jak zlepšit vyhledání odpovědi pro specializované domény. Jedná se o originální řešení, které nebylo zatím publikováno.

Formální a jazyková úroveň, rozsah práce — A-výborně

Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.

Práce je kvalitně zpracována, je třeba ocenit, že je napsána v anglickém jazyce. Práce je vhodně doplněna výstižnými obrázky, grafy. Výsledky jsou zpracovány ve formě tabulek. Rozsah práce je adekvátní.

Výběr zdrojů, korektnost citací — A-výborně

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.

Student přistupoval k řešení velmi aktivně a zodpovědně. Se studentem jsem pracoval skoro dva roky a za tu dobu udělal velký pokrok, který se odrazil v kvalitě předkládané práce. Student se v problematice velmi dobře orientuje a využil nejnovější odborné zdroje pro danou tématiku. Text správně referuje použité zdroje a uvádí vlastní výsledky. Seznam citací je v pořádku.

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.

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

POSUDEK VEDOUCÍHO ZÁVĚREČNÉ PRÁCE

Předkládaná práce je příspěvkem k řešení jedno z problémů, který se vyskytuje v řadě aplikací nejen v odpovídání dotazů. Dosažený výsledek má vysokou úroveň s kvalita je navíc podtržena kvalitními experimentálními výsledky.

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.

Otázka: Kde plánujete algoritmy nasadit, případně publikovat zdrojové kódy do open-source?

Vzhledem k šíři celého projektu a odborné úrovni výsledného řešení, předloženou závěrečnou práci hodnotím klasifikačním stupněm A-výborně.

Datum: 18.6.2018

Podpis:

ČVUT
ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

POSUDEK OPONENTA ZÁVĚREČNÉ PRÁCE

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	**Porozumění nestrukturovanému textu a automatické odpovídání na dotazy**
Jméno autora:	**Bc. Jakub Konrád**
Typ práce:	diplomová
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra řídící techniky
Oponent práce:	Ing. Jan Pichl
Pracoviště oponenta práce:	Katedra kybernetiky

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání — **průměrně náročné**

Hodnocení náročnosti zadání závěrečné práce.

Zadání hodnotím jako průměrně náročné, kdy student sice nevyvíjí vlastní algoritmus pro řešení problematiky, ale musí se velice podrobně seznámit s existujícím řešením. To vyžaduje podrobnou analýzu publikace a zdrojového kódu zkoumaného systému, který nemusí být kvalitně zdokumentován.

Splnění zadání — **splněno**

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.

Splněno bez výhrad.

Zvolený postup řešení — **vynikající**

Posuďte, zda student zvolil správný postup nebo metody řešení.

Student ve své práci věnuje pozornost SQuAD datasetu, který je v dnešní době jedním z nepopulárnějších pro měření výsledků systémů, které se zabývají odpovídáním na otázky v přirozeném jazyce. Vybraný systém DrQA sice není systém, který dosahuje nejlepších výsledků na zvoleném datasetu, ale je to jeden z mála systémů (umístěných na předních pozicích) využívající jeden model nikoliv kombinaci několika modelů.

Odborná úroveň — **A - výborně**

Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.

Práce je na velmi dobré odborné úrovni. Student velice srozumitelně popisuje problematiku neuronových sítí, která je klíčová pro vybraný systém. Na základě vysvětlených obecných principů ANN včetně RNN student podrobně a srozumitelně popisuje princip netriviálního systému DrQA.

Formální a jazyková úroveň, rozsah práce — **A - výborně**

Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.

Po jazykové stránce je práce na vysoké úrovni, kvalita angličtiny je výborná, pouze drobné překlepy. Po formální stránce vykazuje práce pouze drobné nedostatky jako je špatné použití úvodních uvozovek či umístění popisku tabulky pod tabulkou.

Výběr zdrojů, korektnost citací — **A - výborně**

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.

ČVUT
ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

POSUDEK OPONENTA ZÁVĚREČNÉ PRÁCE

Student vzorně cituje veškerá prezentovaná fakta. Při základním popisu správně vybírá původní publikace (např. popis LSTM sítí). V kapitole popisující samotný QA systém jsou fakta podložena aktuálními a relevantními zdroji. Seznam literatury je v souladu s normou.

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.

(Popsáno v celkovém 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.

Student se ve své práci zaměřuje na adaptaci úlohy odpovídání na faktické otázky pro konkrétní doménu, což je užitečné v případech, kdy je daná doména předem známa (např. je určena z kontextu dialogu).

Práce začíná teoretickým popisem neuronových sítí, který je nezbytný po celkové porozumění vybranému systému. Následuje popis vybraného velice populárního datasetu SQuAD a podrobný popis samotného QA systému DrQA. Před provedením samotných experimentů student správně analyzuje druhy chyb, kterých se systém dopouští. V následujících kapitolách popisuje zvolený experiment, který se skládá z vybraných domén. Výsledky experimentů, jakkoliv se zdají být očekávané, jsou popsány srozumitelně a ilustrovány řadou grafů. Zde oceňuji srovnání zastoupení typů chyb s původním datasetem.

Celkově hodnotím práci jako vynikající.

Otázky:

1. Neměl by v tabulce 6.1 sloupec popisující zastoupení témat v trénovací sadě také sčítat do 100 procent?
2. Proč byly vybrány osoby namísto organizací jako testované téma, když organizace jsou v datech více zastoupené (podle tabulky 6.1).
3. I když se zdá, že chyby v kategorii „nesmyslných odpovědí" spolu nijak nesouvisí, nemáte alespoň pro některé z nich intuitivní vysvětlení?
4. Plánujete systém použít v reálné aplikaci? Zmiňoval jste dialogový systém. Ve kterých částech (doménách) dialogu vidíte využití?

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **A - výborně.**

Datum: 25.1.2018 Podpis: