

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Deep Learning for Relational Databases
Jméno autora:	Peleška Jakub
Typ práce:	diplomová
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra počítačů
Oponent práce:	Ing. Jan Drchal, PhD.
Pracoviště oponenta práce:	AIC

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	náročnější
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Problematika zpracování databází pomocí metod hlubokého učení je do značné míry neprozkoumaná. Je třeba detailně nastudovat náročnější témata, jakými jsou např. grafové neuronové sítě. Řešení je implementačně náročné, navíc je třeba řešit infrastrukturu pro uložení a přístup k testovaným databázím.	

Splnění zadání	splněno
<i>Posudte, 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 je splněno bez výhrad. Rozsah práce je značný. Bylo dosaženo velmi zajímavých, publikovatelných výsledků.	

Zvolený postup řešení	vynikající
<i>Posudte, zda student zvolil správný postup nebo metody řešení.</i>	
Se zvoleným postupem řešení se zcela ztotožňuji. Návrh architektury DatabaseTransformer je dobře zdůvodněn.	

Odborná úroveň	A - výborně
<i>Posudte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
Z odborného hlediska je práce na vysoké úrovni. Je zřejmé, že se student v problematice výborně orientuje. Oceňuji solidní srovnání s metodami state-of-the-art a dobře provedenou „ablation study“ testující dílčí komponenty modelu.	

Formální a jazyková úroveň, rozsah práce	C - dobře
<i>Posudte správnost používání formálních zápisů obsažených v práci. Posudte typografickou a jazykovou stránku.</i>	
Text je psán výbornou angličtinou s minimem překlepů. Text je dobře strukturovaný a čtivý, mám k němu však několik méně důležitých námitek: <ul style="list-style-type: none"> Některá témata, na kterých je vaše řešení založeno, by bylo vhodnější lépe vysvětlit. Příkladem může být metoda HGSampling, případně sekce 4.3.3. V sekci 4.1.2 postrádám ukázkou extrahovaného schématu. Čitelnosti by obecně pomohlo více příkladů. Sekce 4.1.4 Graph Construction by si zasloužila pseudokód nebo diagram. Formální zápis je kvalitní, jen někdy chybí rozsahy proměnných. 	

Výběr zdrojů, korektnost citací	A - výborně
<i>Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce. Charakterizujte výběr pramenů. Posudte, 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.</i>	
Zdroje jsou citovány správně. Rozsah referencí je značný.	

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.

Práce je s úpravami publikovatelná. Dosažené výsledky jsou velmi nadějně a mají velký aplikační potenciál.

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.

K práci mám následující dotazy:

- 1) Máte představu o tom, jak bude vaše metoda škálovat na větší databáze? Jak škálování ovlivňuje počet relací (tabulek), jejich velikost, počet integritních omezení a další parametry databáze?
- 2) V textu odkazujícím se k tabulce 5.6 píšete „*In general, the blueprint instances outperformed the selection of the methods with a small number of exceptions where the propositionalization shined.*“ Nemáte představu o tom, jaké jsou vlastnosti těchto databází, pro které je propositionalizace výhodná?

Drobné výtky k textu (viz výše) výrazně převažuje celková kvalita i rozsah práce. Předloženou závěrečnou práci hodnotím klasifikačním stupněm **A - výborně**.

Datum: 11.6.2024

Podpis: