

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Sum-Product-Set Networks for Density Learning of Tree-Structured Data
Jméno autora:	Martin Rektoris
Typ práce:	díplomová
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	katedra kybernetiky
Vedoucí práce:	doc. Ing. Václav Šmídl, Ph.D.
Pracoviště vedoucího práce:	katedra počítačů

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	mimořádně náročné
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Práce se zabývá modelováním hustoty pravděpodobnosti netypickou metodou na netypickém datovém prostoru. Student se tedy musel seznámit s velkým množstvím nové látky a teorie.	

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 bylo splněno. V zadání se vyskytují požadavky na experimenty s jednoduchými daty, autor předkládá experimentální studie již na velkém souboru standardních benchmarkových dat. Tato změna byla motivována dokončením konferenčního příspěvku na téma diplomové práce.	

Aktivita a samostatnost při zpracování práce	A - výborně
<i>Posudte, zda byl student během řešení aktivní, zda dodržoval dohodnuté termíny, jestli své řešení průběžně konzultoval a zda byl na konzultace dostatečně připraven. Posudte schopnost studenta samostatně tvůrčí práce.</i>	
Student pracoval velmi samostatně a velmi iniciativně. Objem vykonané práce je vyšší než je uvedeno v textové části.	

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>	
Autor se důkladně seznámil jak s teorií sum-product sítí tak s teorií bodových procesů. Kombinace obou teorií je bezchybná a experimentální studie odpovídá standardům v oboru.	

Formální a jazyková úroveň, rozsah práce	A - výborně
<i>Posudte správnost používání formálních zápisů obsažených v práci. Posudte typografickou a jazykovou stránku.</i>	
Práce je napsaná velmi dobrou angličtinou, s minimem překlepů a stylových chyb. Menší nedostatky jsou způsobeny nedostatkem času před odevzdáním.	

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>	

Použitá literatura je vyčerpávající a správně citovaná. Autor aktivně vyhledával literaturu a použité citace pečlivě studoval.

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 významně přispěla k dokončení článku na konferenci ICRL (A*), který je nyní v recenzním řízení a obdržel dobré hodnocení.

- Sum-Product-Set Networks: Deep Tractable Models for Tree-Structured Graphs, M. Papež, M. Rektoris, V. Šmídl, T. Pevný, ICLR 2024

III. CELKOVÉ HODNOCENÍ A NÁVRH KLASIFIKACE

Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení.

Práce spočívala v nastudování dvou nestandardních teorií a aplikování jejich kombinace na netypická data. Auto pracoval velmi pilně a samostatně. Výsledkem je výborná práce, která přispěla k zaslání článku na dané téma na A konferenci.*

Předloženou závěrečnou práci hodnotím klasifikačním stupněm A - výborně.

Datum: 14.1.2024

Podpis:

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Sum-Product-Set Networks for Density Learning of Tree-Structured Data
Jméno autora:	Martin Rektoris
Typ práce:	diplomová
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra řídicí techniky
Oponent práce:	Jiří Vomlel
Pracoviště oponenta práce:	ÚTIA AV ČR

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	náročnější
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Zadání práce vychází z nejnovějších výsledků a trendů v oblasti modelování datových struktur. Z tohoto pohledu považuji zadání za náročnější, protože řešení práce vyžaduje pečlivé seznámení s nejnovějšími výsledky ve studované oblasti.	

Splnění zadání	splněno s menšími výhradami
<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>	
Prvním cílem uvedeným v zadání byl přehled nejnovějších poznatků o <i>sum-product networks</i> (SPNs) ilustrovaný na příkladech a doplněný o shrnutí výhod a nevýhod tohoto modelu. <i>Sum-product networks</i> jsou v práci popsány dostatečně detailně, ale přehledné shrnutí výhod a nevýhod tohoto modelu jsem v práci nenalezl. Dalším cílem bylo představení modelu náhodných konečných množin, jeho vlastností a využití, který měl být ilustrovaný na jednoduchých příkladech. Pravděpodobnostní model náhodných konečných množin je v práci definován a jsou popsány jeho základní vlastnosti. Popisu využití a vysvětlení na příkladech by dle mého názoru mohlo být přehlednější a pečlivější. Třetí cíl, kterým bylo rozšíření <i>sum-product networks</i> o další uzel modelující náhodnou konečnou množinu a prozkoumání vlastností tohoto modelu, byl splněn a tento nový model je v práci představen. Stejně tak byl splněn poslední cíl, kterým bylo ověření vlastností nového modelu na souboru referenčních příkladů stromově strukturovaných dat a porovnání s konkurenčními architekturami, jako jsou hluboké neuronové sítě.	

Zvolený postup řešení	správný
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
Hlavní argument pro výzkum této problematiky vychází z úvahy, že k vývoji konkurenceschopné umělé inteligence založené na strojovém učení bychom jí měli v ideálním případě poskytnout stejné (strukturované) vstupy jako lidem. Příkladem formátu, který tento požadavek splňuje, je v práci využitý formát JSON. Strukturované formáty jsou využity v nově navržených modelech <i>sum-product set networks</i> (SPSNs).	

Odborná úroveň	B - velmi dobře
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
Autor v práci prokazuje velmi dobrou odbornou znalost studované problematiky. V některých částech textu nejsou použité odborné termíny definovány nebo nedostatečně vysvětleny. Např. v Definiční 6 se objevuje V a V namísto V a E. V úvodu sekce 3.2.1 je na straně 22 zaveden pojem „bag“ pro jehož zavedení je použita množina $F(X)$, která je ovšem definována až na straně 25. Pozorování je definováno jako množina, ale není řečeno, co jsou prvky této množiny. Na straně 25 se mluví o „random proces“, ten ale nebyl nikde definován ani popsán. Na straně 36 se objevují dva nadpisy sekcí, které ale nemají žádný obsah.	

Formální a jazyková úroveň, rozsah práce

C - dobře

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

Práce je psána v anglickém jazyce na dobré jazykové úrovni, v některých místech jsou drobné jazykové nedostatky. Např. „S is constructed such by recursively traversing“, „each element is transform its corespodning function“, „the tree can traversed using“, apod. Práce je spíše menšího rozsahu, dle mého názoru to ale nepředstavuje problém, protože důležitější než rozsah je obsah.

Výběr zdrojů, korektnost citací

A - výborně

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.

Výběr zdrojů a jejich citace je na výborné úrovni a splňuje nejvyšší standardy vědecké práce. Bohužel odkazovaná data z experimentů nebyla v době přípravy posudku dostupná na adrese uvedené v diplomové práci. <https://relational.fit.cvut.cz/>.

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 neobsahuje zásadně nové teoretické výsledky. Hlavním původním přínosem práce je návrh zobecnění *sum-product networks* (SPNs) na *sum-product set networks* (SPSNs). Toto rozšíření je implementováno a poměrně rozsáhle otestováno. Myslím, že tento výstup je velmi významný a práce díky tomu nepochybně splňuje požadavky kladené na diplomovou práci. Drobnější slabinou jsou některá chybějící vysvětlení a nepřesná vyjádření. Textu by bývala ještě prospěla pečlivá finální kontrola.

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.

Navzdory výše uvedeným (spíše drobnějším) nedostatkům považuji práci za velmi kvalitní. Obzvláště oceňuji rozsáhlé a detailní experimenty. Práce může dobře posloužit jako východisko pro další výzkum v této moderní a atraktivní oblasti výzkumu.

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

Datum: 2.2.2024

Podpis: