

Posudek oponenta závěrečné práce

České vysoké učení technické v Praze

Fakulta informačních technologií

Student: Karel Rymeš
Oponent práce: Ing. Daniel Vašata, Ph.D.
Název práce: Neural networks as anomaly detectors
Obor: Teoretická informatika

Datum vytvoření: 30. 1. 2017

Hodnotící kritérium: 1. Náročnost a další komentář k zadání	Způsob hodnocení - následující škálou 1 až 5: <u>1=mimořádně náročné zadání,</u> 2=náročnější zadání, 3=průměrně náročné zadání, 4=lehčí, ale ještě dostatečně náročné zadání, 5=nedostatečně náročné zadání
Popis kritéria: Podrobněji charakterizujte diplomovou (bakalářskou) práci a její případné návaznosti na předchozí nebo běžící projekty. Dále posuďte, čím je zadání této ZP náročné. (U obtížnější ZP lze dále tolerovat některé nedostatky, které by u ZP standardní obtížnosti tolerovány nebyly; a naopak u jednoduché ZP mohou být zjištěné nedostatky hodnoceny přísněji.) Komentář: Zadání považuji za mimořádně náročné, protože je třeba prostudovat poměrně obtížné matematické koncepty kombinující prvky teorie pravděpodobnosti, statistiky a optimalizace.	
Hodnotící kritérium: 2. Splnění zadání	Způsob hodnocení - následující škálou 1 až 4: 1=zadání splněno, <u>2=zadání splněno s menšími výhradami,</u> 3=zadání splněno s většími výhradami, 4=zadání nesplněno
Popis kritéria: Posuďte, zda předložená ZP splňuje zadání. V komentáři uveďte body zadání, které nebyly zcela splněny, případně rozšíření ZP oproti původnímu zadání. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků. Komentář: Rešeršní část práce je na velmi dobré úrovni. Drobnou výhradu mám k experimentální části, kde bylo cílem analyzovat vliv hyper-parametrů u vybraného typu neuronových sítí na schopnost generalizace, efektivitu učení a kvalitu detekce anomálií. Tato analýza je ovšem velmi strohá a skládá se v podstatě pouze ze stručného popisu výsledků a konstatování, že vliv hyper-parametrů nebyl vzhledem k variabilitě výsledků prezentovatelný.	
Hodnotící kritérium: 3. Rozsah písemné zprávy	Způsob hodnocení - následující škálou 1 až 4: <u>1=splňuje požadavky,</u> 2=splňuje požadavky s menšími výhradami, 3=splňuje požadavky s většími výhradami, 4=nesplňuje požadavky
Popis kritéria: Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Komentář: Zcela v pořádku.	
Hodnotící kritérium: 4. Věcná a logická úroveň práce	Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F): 90 (A)
Popis kritéria: Posuďte, zda předložená ZP je po věcné stránce v pořádku, případně vyskytují-li se v práci věcné chyby nebo nepřesnosti. Zhodnoťte dále logickou strukturu ZP, návaznosti jednotlivých kapitol a pochopitelnost textu pro čtenáře. Komentář: Práce je přehledná a srozumitelná. Po věcné stránce obsahuje pouze drobné nedostatky. Například by v práci mohl být podrobněji diskutován algoritmus zpětného přenosu (back-propagation). Především jeho zápis na straně 32 není úplně srozumitelný. Ve vztazích 3.17 a 3.18 by levé strany neměly být označovány parciálními derivacemi nebo by měla být zdefinována funkce J a ukázáno, že výrazy na pravé straně skutečně odpovídají jejím derivacím. Dále by bylo v experimentální části v kapitole 5 vhodné podrobněji rozvést použité metody. V práci například není uvedeno jaká aktualizace algoritmu zpětného přenosu (back-propagation update) ze str. 31 byla použita.	
Hodnotící kritérium: 5. Formální úroveň práce	Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F): 75 (C)
Popis kritéria: Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 14/2015, článek 3.	

Komentář:

Práce je jazykově na velmi slušné úrovni. Z typografického pohledu se v ní však vyskytuje řada nedokonalostí. Mezi ně patří především začínání řádků čárkami, které následují za vzorci (strana 41), dále používání skloněného fontu pro operátory v matematických výrazech (např. \max v (3.10), ME v (1.5), apod.) a také odkazování na číslované obrázky a kapitoly s použitím malého písma na začátku (např. na konci strany 27).

Hodnotící kritérium:

Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):

6. Práce se zdroji

95 (A)

Popis kritéria:

Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení ZP. Charakterizujte výběr studijních pramenů. Posuďte, zda student využil všechny relevantní zdroje nebo zda se pokoušel řešit již vyřešené problémy. 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.

Komentář:

Práce se zdroji je na vynikající úrovni.

Hodnotící kritérium:

Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):

7. Hodnocení výsledků, publikační výstupy a ocenění

85 (B)

Popis kritéria:

Vyjádřete se k úrovni dosažených hlavních výsledků ZP, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, apod. Případně také zhodnoťte, zda software nebo zdrojové texty, které nevytvořil sám student, byly v ZP použity v souladu s licenčními podmínkami a autorským právem. Popište případnou publikační činnost a získaná ocenění související s řešením této ZP.

Komentář:

Teoretická část práce je velmi pěkně zpracovaná. U experimentální části bych si dovedl představit mnohem podrobnější popis získaných výsledků především co se týče závislosti efektivity učení a kvality detekce anomálií na různých volbách hyper-parametrů. Jediné porovnání tohoto typu je uvedeno v tabulkách 5.2 a 5.3, kde jsou ovšem pouze dílčí výsledky, které podle mého názoru zdaleka nepředstavují splnění této části zadání.

Hodnotící kritérium:

Způsob hodnocení - nehodnotí se

8. Komentář o využitelnosti výsledků

Popis kritéria:

Uvedte hlavní výsledky ZP rozšiřující již publikované známé výsledky a/nebo přinášející zcela nové poznatky. Uvedte možnosti využití výsledků ZP v praxi.

Komentář:

Teoretická část práce může být využita jako úvodní přehled teorie neuronových sítí se zaměřením na detekci anomálií. Experimentální část práce je využitelná mnohem méně a spíše jen jako vodítko ke konkrétnější volbě modelu a některých jeho hyper-parametrů.

Hodnotící kritérium:

Způsob hodnocení - nehodnotí se

9. Otázky k obhajobě

Popis kritéria:

Uvedte případné dotazy, které by měl student zodpovědět při obhajobě ZP před komisí (body oddělte odrážkami).

Otázky:

Na str. 50 píšete, že závislost kvality detekce anomálií na volbě hyper-parametrů nebyla použitelná. Lze alespoň na obecné úrovni, s přihlédnutím k výpočetní složitosti učicí fáze, říci, zda je u stacked autoencodéru lepší zvyšovat počet neuronů v rámci jedné vrstvy nebo zvyšovat počet vrstev?

Hodnotící kritérium:

Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):

10. Celkové hodnocení

89 (B)

Popis kritéria:

Shrňte stránky ZP studenta, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení **ne** musí být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích 1 až 9.

Text hodnocení:

Vzhledem k obtížnosti zadání nepovažuji výše uvedené nedostatky za příliš závažné a celkově hodnotím stupněm B.

Podpis oponenta práce: