

Posudek oponenta závěrečné práce

České vysoké učení technické v Praze

Fakulta informačních technologií

Student: Matěj Nikl
Oponent práce: doc. RNDr. Ing. Marcel Jiřina, Ph.D.
Název práce: Solving text-based CAPTCHA images with neural networks
Obor: Teoretická informatika

Datum vytvoření: 15. 6. 2017

Hodnotící kritérium: 1. Náročnost a další komentář k zadání	Způsob hodnocení - následující škálou 1 až 5: 1=mimořádně náročné zadání, 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 práce považuji co do náročnosti za standardní a adekvátní studovanému oboru.	
Hodnotící kritérium: 2. Splnění zadání	Způsob hodnocení - následující škálou 1 až 4: 1=zadání splněno, 2=zadání splněno s menšími výhradami, 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: Hlavní cíl i všechny dílčí cíle zadání byly splněny. Určité pochybnosti mám ohledně rozsahu rešerše, která je dle mého názoru nedostačující a dále ohledně vyhodnocení výsledků.	
Hodnotící kritérium: 3. Rozsah písemné zprávy	Způsob hodnocení - následující škálou 1 až 4: 1=splňuje požadavky, 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: Práce obsahuje cca 33 stran hlavního textu, což je v pořádku, nicméně si dokáží představit, že rešerše bude podrobnější, zejména v kapitolách 3.2 a 3.3, vlastní návrh řešení v kapitole 4 bude rozveden na více než 3 strany a experimenty budou bohatší a lépe popsány.	
Hodnotící kritérium: 4. Věcná a logická úroveň práce	Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F): 85 (B)
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 systematicky a věcně vystavěna. Všechny kapitoly na sebe navazují a tvoří ucelený logický rámec.	
Hodnotící kritérium: 5. Formální úroveň práce	Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F): 85 (B)
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: Po typografické, jazykové i stylistické stránce je práce na přiměřeně dobré úrovni. Některé popisy jsou nepřesné a nejednoznačné a mohou být špatně interpretovány.	
Hodnotící kritérium: 6. Práce se zdroji	Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F): 75 (C)

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 uvádí 37 zdrojů, které jsou v souladu s citačními zvyklostmi a normami. V kapitole 3, zejména pak v podkapitolách 3.2 a 3.3, je uvedena jen jedna relevantní citace v každé kapitole. To považuji za nedostatečné, zejména pokud uvážím, že při vlastním velmi rychlém průzkumu článků, které se zabývají prolamováním CAPTCHA obrázků pomocí neuronových sítí, jsem našel hned několik relevantních publikací, které student neuvádí. Dále v práci na řadě míst chybí odkazy na literaturu. Např. v kap. 5.1 na str. 25 postrádám odkaz na zdroj, který byl použit pro generování experimentálních dat.

Hodnotící kritérium:

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

7. Hodnocení výsledků, publikační výstupy a ocenění

80 (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:

Mám určité pochybnosti o kvalitě provedených experimentů. V kapitole 5.2 autor zmiňuje, že trénovací množina měla 100 000 vzorů, ale nikde jsem nenašel, kolik měla testovací množina a zda výsledky (dosažené přesnosti) prezentované na grafech 5.2, 5.7 a 5.8, resp. v tabulkách 5.1, 5.2 a 5.3, jsou získány na základě testovací množiny.

Hodnotící kritérium:

Způsob hodnocení - nehodnotí se

8. Komentář o využitelnosti výsledků

Popis kritéria:

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

Komentář:

Věřím, že dosažené výsledky jsou použitelné v praxi. Navržené řešení je jednoduché a současně funkční, takže i přes určité pochyby ohledně korektnosti prezentovaných výsledků považuji zvolené řešení za dobré a využitelné v praxi.

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:

- 1) Podrobněji vysvětlete, jakým způsobem byla vyhodnocována kvalita neuronové sítě, tj. schopnost rozpoznat (prolomit) CAPTCHA obrázky.
- 2) U jakých typů CAPTCHA obrázků dochází k největší chybovosti?

Hodnotící kritérium:

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

10. Celkové hodnocení

80 (B)

Popis kritéria:

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

Text hodnocení:

Zvolený přístup k řešení práce je poměrně přímočarý, současně logicky správný a funkční, takže práci považuji přes uvedené výhrady za splněnou.

Podpis oponenta práce: