

Strukturovaný posudek oponenta bakalářské práce

Autor/autorka práce: **Miroslav Liška**

Název práce: **Automatická anotace obrázku**

Obsah práce:

Vynikající logická struktura, nadprůměrný obsah i rozsah;
Velmi dobrá logická struktura, odpovídající obsah i rozsah;
Vyhovující logická struktura, obsah i rozsah;
Nevyhovující

Komentář: Cílem práce je na základě prostudované literatury navrhnout a implementovat systém pro automatickou anotaci obrázků. Vytvořený systém bude dále otestován na dodaných datových kolekcích.

Kvalita řešení a dosažených výsledků:

Vynikající; Velmi dobrá; Vyhovující; Nevyhovující

Formální úroveň:

Vynikající; Velmi dobrá; Vyhovující; Nevyhovující

Komentář: Práce obsahuje občasné překlady a nepřesné formulace. Konkrétní připomínky uvádím níže. Rastrové obrázky (pokud se nejedná o fotografie) by bylo vhodné nahradit vektorovými. První číslovaná kapitola má být úvod. Zkratky (např. LBP) by měly definovány při prvním použití. V případě, kde je to možné (např. Cluster), doporučuji používat české ekvivalenty původních anglických termínů. Obrázek 3.3. není obrázek, ale rovnice. Pozitivně hodnotím zejména použití textového procesoru LaTeX a přehlednou uživatelskou příručku v závěru práce. Přiložené CD neobsahuje readme soubor v kořeni, tento soubor je ale v adresáři ProgramoveVybaveni.

Práce s literaturou:

Vynikající; Velmi dobrá; Vyhovující; Nevyhovující

Splnění zadání:

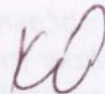
Splněno bez výhrad; Splněno s menšími výhradami; Splněno s většími výhradami;
Nesplněno

Připomínky/dotazy k práci:

1. Neuronové sítě (i klasifikační stromy) klasifikují prvky do několika tříd. Vysvětlíte, proč je v sekce 3.1.2. uvádíte jako binární klasifikaci?
2. V sekci 7 popisujete datové kolekce. Chybí zde ale většinou statistické informace o datech, např. počet kategorií, klíč. slov apod. Objasněte prosím.
3. V úvodu kap. 10 píšete, že všechny parametry byly voleny „čistě experimentálně“. Vysvětlíte prosím jak?
4. V tabulce 10.1. používáte metriku precision (přesnost). Jako oponent nevím, zda jste dělal klasifikaci do více tříd, či nikoli (není možno vyčíst ani z popisu dat). V případě, že klasifikujete to jedné třídy, je vhodné použít metriku accuracy,

- pokud jste klasifikoval do více tříd, chybí mně zde úplnost a F-míra. Objasněte prosím.
5. V sekci 10.2. uvádíte pouze průměrný výsledek. Vzhledem k pravděpodobnému nevyvážení jednotlivých tříd se mi nejeví jako vhodná metrika. Vysvětlete, jak jste přesně počítal? Proč neuvádíte i výsledky klasifikace jednotlivých tříd? Můžete je ukázat? Dále mně zde chybí opět F-míra.
 6. V sekcích 10.3.1 a 10.3.2 srovnáváte dobu běhu předzpracování a anotování. Z mého pohledu není tato hodnota příliš důležitá. Navíc, pokud ji uvádíte, je třeba také uvést, na jakém počítači probíhal výpočet.
 7. V sekci 10.4 srovnáváte Vaše výsledky s výsledky z literatury. Vaše výsledky jsou v tab. 10.2 (případně 10.6), zatímco referenční jsou uvedeny v tab. 10.5 (na jiných stranách) . Pro přehlednost by bylo vhodné dát výsledky do jedné tabulky. Navíc z tab. 10.5 není poznat na jakých datech a jakým způsobem se testovalo. Tabulka 10.6 má zřejmě chybný popisek. Objasněte prosím.
 8. Pro překlad programu je nutné mít nainstalované Visual studio. Postrádám makefile soubor pro překlad Vašeho programu. Připravte prosím.

Navrhuji hodnocení známkou **velmi dobře** a práci doporučuji k obhajobě.



V Plzni 24.5.2015

doc. Ing. Pavel Král, Ph.D.