



Posudek oponenta bakalářské práce

Filip Hácha

Metody strojové klasifikace pro výběr optimálního estimátoru křivosti

Předložená práce navazuje na výzkumnou aktivitu centra, která se zabývá výpočtem křivosti trojúhelníkových sítí.

Text práce je vcelku logicky členěn do 12 kapitol. Po krátké úvodní kapitole je zařazena kapitola věnovaná diferenciální geometrii, která velmi stručně seznamuje čtenáře s problematikou výpočtu křivosti u spojitých křivek a ploch. V další kapitole pak bakalant opět velmi zjednodušeně popisuje problematiku výpočtu křivosti pro trojúhelníkové sítě. 4. kapitola ve stručnosti shrnuje metody strojového učení a typy úloh, které lze pomocí strojového učení řešit. V kapitole 5 je čtenář možná až příliš stručně seznámen s formulací samotné úlohy. Bylo by zde asi na místě konkrétněji představit jaké typy vlastností se u trojúhelníkových sítí zkoumají, jaké typy sítí se v testu vyskytují a proč není vhodné pro všechny sítě využít stejný estimátor, mezi jakými estimátory se vlastně bude vybírat a jaké jsou jejich vlastnosti. Práce je sice postavena obecně a je možné v budoucnu libovolně přidávat nové estimátory, nezasvěcený čtenář ale pravděpodobně nebude mít ani přibližnou představu. Textu by v tomto místě určitě pomohlo i několik vzorových obrázků. Kapitola 6 se věnuje normalizaci jednotlivých vlastností. Na tomto místě by možná stálo za to zmínit, že pro některé metody může být tato normalizace klíčová (je zmíněná metoda shlukování, u které jsou rozsahy jednotlivých vlastností zásadní, zároveň ale nejsem přesvědčen, že normalizace na jednotkový interval je správná – výzkum v tomto směru je však zcela jistě nad rámec bakalářské práce). Kapitoly 7 a 8 stručně popisují jednotlivé klasifikační metody a způsoby vyhodnocení dosažených výsledků. Nejděší, nikoliv však stěžejní, kapitolou je kapitola věnovaná implementačním detailům. Osobně bych uvítal, kdyby byly detailněji zpracované teoretičtější části práce a tato kapitola byla výrazně kratší. Kapitola 10 pak slouží jako stručná uživatelská příručka a mohla být spíše součástí příloh. Bližší rozbor by si naopak zasloužily výsledky uvedené v kapitole 11, nejsem si ale jist, zda je v silách studenta bakalářského studia takový rozbor udělat. Alespoň částečné zhodnocení výsledků je pak provedeno v závěrečné kapitole.

Celkově mi text práce připadá nevyvážený, je ale nutné podotknout, že zadání práce je na výrazně vyšší úrovni než jsou běžná bakalářská zadání. Pro bakalanta je tak nesmírně obtížné vybrat z množství nových znalostí, které musel pro práci zpracovat, ty důležité.

Dodané programové vybavení je na takto experimentální typ zadání až překvapivě dobře navrženo. Jako drobný nedostatek bych viděl absenci dokumentace u velké části zdrojových kódů, což by mohlo zkomplikovat budoucí navazující práce.

I přes drobné výtky ke zpracování je nutné říct, že bakalant odvedl velký kus netriviální vědecké práce a kvalita práce přesahuje běžné bakalářské práce. Bakalant zároveň prokázal, že je schopen sestudovat netriviální metody z různých oblastí a implementovat vhodné řešení pro experimentální ověření postupu nad netriviálními daty. Z práce není patrné, kolik času zabralo samotné měření, ale z rozsahu dat a počtu měřených metod bych předpokládal, že samotné experimenty byly časově velmi náročné. Práci doporučuji k obhajobě a hodnotím stupněm

„výborně“

Zároveň doporučuji zvážit, zda danou práci nenavrhnout na některou z cen.

V Plzni 15. května 2019

Ing. Petr Vaněček, Ph.D.
(oponent BP)

Doplňující otázky:

- Jako nejlepší metoda v tuto chvíli vychází rozhodovací strom stejně jako v původní práci, kde byl metaestimátor navržen ručně. Shodují se estimátory vybrané metodami strojového učení s těmi z původního návrhu nebo se jedná o jinou podmnožinu estimátorů?
- Pokud dobře rozumím výsledkům v příloze B, nejlepší estimátory dosahují úspěšnosti kolem 30%, což mi připadá velmi málo. V jakých hodnotách se rozdíly u jednotlivých estimátorů vlastně pohybují?