



# Metody strojové klasifikace pro výběr optimálního estimátoru křivosti

Filip Hácha<sup>1</sup>

## 1 Úvod

Tato práce se zabývá zkoumáním několika různých přístupů z oblasti strojového učení pro určení optimálního estimátoru křivosti trojúhelníkových sítí na základě jejich vlastností. V práci jsou prezentovány výsledky dosažené za pomoci různých klasifikačních algoritmů.

Při konstrukci klasifikátorů některých klasifikátorů byly dále využity optimalizační pro nalezení jeho optimální konfigurace, algoritmy pro normalizaci a analýza hlavních komponent pro redukci dimenze vstupních dat.

## 2 Křivost

Křivost je zajímavou vlastností ploch, neboť umožňuje rozpoznání rysů jako jsou údolí, roviny, konvexní, konkávní nebo sedlové tvary, jak ukazují Gatzke a Grimm (2006). Jak poukázali Váša et al. (2017), zatímco výpočet křivosti v bodě na hladkých površích běžně nepředstavuje obtížnou úlohu, odhad křivosti v bodech polygonálních sítí je obtížnější, neboť rozdílná tělesa mohou být reprezentovány stejnou trojúhelníkovou sítí.

Způsobů, jak odhadovat křivost původního tělesa na základě trojúhelníkové sítě, je více a pro různá data nám poskytují různé estimátory odhady s rozdílnou přesností. Výběr optimálního estimátoru pro konkrétní síť tedy můžeme označit za klasifikační úlohu.

## 3 Předchozí práce

Tato práce vychází z výzkumu, který v této oblasti provedli Váša et al. (2016), který se zabýval implementací jednotlivých estimátorů křivosti a ve kterém byl zároveň popsán klasifikátor založený na rozhodovacích stromech s jedním a se dvěma rozhodovacími uzly.

## 4 Metody klasifikace

Vstupními daty úlohy je množina trojúhelníkových sítí, která je parametrizována vektorem různých vlastností a chybou jednotlivých estimátorů na dané síti. Cílem úlohy je přiřadit sítím na základě vektoru sledovaných vlastností optimální estimátor křivosti.

Rozsahy hodnot sledovaných vlastností se od sebe řádově liší, proto je pro některé klasifikační algoritmy nutné provést jejich normalizaci. Prvním implementovaným přístupem je normalizace vlastností, jelikož je tato metoda ale velmi náchylná na odlehlá měření, byla později nahrazena normalizací směrodatnou odchylkou.

Pro některé některé způsoby klasifikace je počet sledovaných vlastností příliš vysoký a jelikož lze předpokládat, že některé z vlastností mohou být navzájem závislé, je možné provést

---

<sup>1</sup> student bakalářského studijního programu Aplikované vědy a informatika, obor Informatika, e-mail: ha-chaf@kiv.zcu.cz

jejich redukci. Pro tento účel byla implementována analýza hlavních komponent, která umožnila např. prohledání větší části prostoru možných řešení při konstrukci rozhodovacího stromu.

Práce porovnává úspěšnost klasifikace trojúhelníkových sítí s využitím klasifikátorů, které byly zkonstruovány za pomoci rozhodovacích stromů, neuronových sítí, algoritmu pro shlukování a lineární  $L^2$  aproximace. Pro konstrukci rozhodovacích stromů s více vrstvami byly vyzkoušeny optimalizační metody simulovaného žíhání a genetické algoritmy, které spolu s redukcí počtu sledovaných vlastností umožnily konstrukci stromů s větším počtem rozhodovacích uzlů

Úspěšnost jednotlivých klasifikátorů byla porovnána stejně jako ve výchozí práci pomocí relativní chyby vypočtené podle vzorce 1.

$$e_{rel} = \frac{e_i - e_{min}}{e_{min}} \quad (1)$$

## 5 Výsledky

Nejlepších výsledků bylo dosaženo zredukováním dimenze vstupních dat analýzou hlavních komponent a následnou konstrukcí binárního rozhodovacího stromu se třemi vrstvami rozhodovacích uzlů za využití optimalizace pomocí genetického algoritmu. Takto sestavený klasifikátor dosáhl relativní chyby  $e_{rel} = 0.59$ , což je zlepšení oproti klasifikátoru, který uvedli Váša et al. (2016), s relativní chybou  $e_{rel} = 0.86$ .

## 6 Závěr

V průběhu práce se podařilo dosáhnout výrazného zlepšení úspěšnosti klasifikace trojúhelníkových sítí oproti výchozímu výzkumu, můžeme však očekávat, že u některých použitých klasifikačních algoritmů nebyl zcela naplněn jejich potenciál a je zde tedy prostor k dalšímu výzkumu.

V budoucí práci je například možné hlouběji prozkoumat možnost použití neuronových sítí a také sofistikovanější metody pro konstrukci rozhodovacích stromů, jako například využití TDIDT (Top down induction of decision trees) algoritmů.

## Poděkování

Děkuji panu doc. Ing. Liborovi Vášovi, Ph.D. za ochotu při vedení bakalářské práce a rady s jejím vypracováním. Dále děkuji panu doc. Ing. Pavlovi Královi, Ph.D. za poskytnutí konzultací k problematice neuronových sítí.

## Literatura

- D. Gatzke, T. a Grimm, C. (2006) *Estimating Curvature on Triangular Meshes* International Journal of Shape Modeling, Volume 12, pp. 1-28.
- Váša, L. a Kühnert, T. a Brunnett G. (2017) *Multivariate analysis of curvature estimators* Computer-Aided Design and Applications, Volume 14, pp. 58-69.
- Váša, L. a Vaneček, P. a Prantl, M. a Skorkovská, V. a Martínek, P. a Kolingerová, I. (2016) *Mesh Statistics for Robust Curvature Estimation* Eurographics Symposium on Geometry Processing, Volumen 35, pp. 271-280.