

Aleh Masaila: Regresní stromy
(Posudek práce podané v dubnu 2013)

Předložená práce je věnována postupu, který je alternativou regresních modelů při hledání aproximace závislosti jedné proměnné na jedné či několika vysvětlujících proměnných. Jde o regresní či rozhodovací stromy v závislosti na charakteru vysvětlované proměnné. Jedná se skutečně o aproximaci, na rozdíl od regresního modelu se nepředpokládá žádný konkrétní tvar závislosti, v níž bychom pouze nahrazovali neznámé parametry příslušnými odhady.

Autor práce shromáždil nemalé množství materiálu, zvolené metody postupně vysvětlil a na poměrně rozsáhlých datových souborech dostupných v knihovně prostředí R ukázal chování vysvětlovaných metod včetně porovnání s klasickým regresním modelem.

Provedení práce, i po přepracování, je bohužel značně nedokonalé. I když vím, že čeština není mateřským jazykem diplomanta, přesto považuji jazykovou úroveň za zcela nedostatečnou. Text místy postrádá matematickou přesnost a jednoznačnost (např. formulace věty 4.2 na str. 34). Vyskytují se tu nedefinované pojmy (patro stromu na str. 40). Slovo *významnost* má ve statistice svůj ustálený význam, ale vysvětlení pojmu statistická významnost na str. 16 mu neodpovídá. Také interpretace testu hypotézy $\beta_j = 0$ v modelu mnohonásobné lineární regrese je poněkud složitější, než je uvedeno na str. 5. Nejsem si jist významem zápisu $H[+(X_j - s)]$ na str. 20.

Spíše formálních nedostatků nebo neobratností lze uvést celou řadu. Např. vyjádření *způsob výpočtu* nepovažuji za synonymum pro definici kritéria umožňujícího porovnat různé metody. Překvapivý je také výskyt symbolů y_i a současně y_{ki} v (4.7) na str. 28.

Domnívám se, že nemá smysl vyjmenovávat zde další nedopatření, která se v textu vyskytují. Naštěstí diplomant projevil v 5. kapitole slušnou schopnost použít a samostatně porovnat mezi sebou metody, kterým je celá práce věnována. Navíc dlužno konstatovat, že zvolené téma je obtížné. Současně věřím, že při obhajobě uvede diplomant na pravou míru zde vytykané nedostatky. Proto **navrhuji, aby předložená práce byla uznána jako práce diplomová.**

V Praze 13. 5. 2013

Karel Zvára