

# Posudek

vedoucího  oponenta  
 diplomové  bakalářské práce

Autor: **Samuel Říha**

Název práce: **Maximalizace Giniho koeficientu v binární logistické regresi**

Jméno vedoucího: **Mgr. Tomáš Hanzák**

## Matematická úroveň:

vynikající  velmi dobrá  průměrná  podprůměrná  nevyhovující

## Grafická, jazyková a formální úroveň:

vynikající  velmi dobrá  průměrná  podprůměrná  nevyhovující

## Výsledky:

originální  původní i převzaté  netriviální kompilace  citované z literatury  opsané

## Použité metody:

nestandardní  standardní  obojí

## Aplikovatelnost:

přínos pro teorii  přínos pro praxi  přínos pro praxi i teorii  bez přínosu  
 nedovedu posoudit

## Věcné chyby:

téměř žádné  vzhledem k rozsahu a tématu přiměřeně  méně podstatné četné  závažné

## Tiskové chyby:

téměř žádné  vzhledem k rozsahu přiměřený počet  četné

## Celková úroveň práce:

vynikající  velmi dobrá  průměrná  podprůměrná  nevyhovující

**Práci doporučuji uznat jako bakalářskou. Návrh klasifikace přikládám na zvláštním papíru.**

## Připomínky a vyjádření vedoucího:

Téma práce nebylo vůbec jednoduché a je třeba u studenta ocenit jeho velkou samostatnost, schopnost studia odborné literatury (Kap. 2), kreativitu a odvahu pustit se do náročného matematického odvozování (Kap. 4). Práce obsahuje přehledný a účelný popis modelu binární logistické regrese, odhadových technik (ML a OLS) a ukazatelů diversifikace (Gini, K-S,  $R^2$ ).

Teoretická Kap. 2 o ztrátových funkcích je zajímavá, avšak její využití pro zadaný cíl práce zůstalo jen v základu. Z tří praktických metod navržených v Kap. 4 (jejich prezentace zájemci o implementaci mohla být přívětivější) zůstala bohužel jedna nevyzkoušená a dvě se ukázaly jako v praxi nefunkční. Kap. 5 by měla být podrobnější a přinášet jasnější praktické závěry. Celkově uspořádání práce mohlo být lepší a měla se více soustředit na zadaný cíl tématu. To je také hlavní důvod horšího celkového hodnocení práce.

Konkrétní připomínky k práci:

1. Matematicko-formulační nepřesnosti: poznámka na str. 2, nesplnitelná definice „dostatečně bohaté množiny funkcí“ (str. 5), „hladké“ funkce (str. 6 a 11).
2. V textu úplně chybějí odkazy na použitou literaturu.
3. Bavíme-li se o dvou iteračních algoritmech, kde oba konvergují ke správné hodnotě (nebo ne?), jak může být rychlejší metoda méně přesná a naopak (str. 15)? Není zde „přesnost“ de facto jen jiným vyjádřením pro rychlost konvergence metody?
4. Na str. 21 autor tvrdí, že diversifikační schopnost modelu je výborná, pokud McFaddenovo  $R^2$  je  $> 0.3$ . Podle čeho tak usuzuje? Značí tedy např.  $R^2 = 0.05$  nutně slabou diversifikační schopnost?
5. Je metoda v Kap. 4.1 speciálním případem metody z Kap. 4.2? Jak jinak by autor volil distribuční funkci  $F$  než normální?
6. Podivný obrázek 5.3 bez dostatečného vysvětlení.

V Praze dne 29. 8. 2012

Mgr. Tomáš Hanzák