

Oponentní posudek k disertační práci

I. Identifikační údaje

Název disertační práce:	Modeling of extensive files of discrete data
Jméno autora:	Ing. Šárka Jozová
Školící pracoviště:	K611
Oponent práce:	Prof. Ing. Václav Matoušek, CSc.
Pracoviště oponenta práce:	Západočeská univerzita v Plzni, Fakulta aplikovaných věd

II. Hodnocení jednotlivých kritérií

Aktuálnost tématu disertační práce	
Téma disertační práce je aktuální – představuje vytvoření modelu pro klasifikaci nejen dopravních dat s redukcí jejich rozměru na bázi shlukování dat a následné konstrukce lokálního modelu. Práce je založena na dvou cílech – vývoj metody pro modelování a predikci rozsáhlých množin diskretních dat a jejím porovnání s dosud existujícími metodami predikce. Oba cíle jsou v práci poměrně přesně vymezeny a jsou stručně shrnuty v odstavci 1.3. Vývoj a funkce metod i zvolená data jsou v práci podrobně popsány a výsledky testování uvedeny. Z toho pohledu je práce originální a pro další rozvoj podobných modelů přínosná. Dosažené výsledky jsou pak ilustrovány v obsáhlých přílohách.	
Hodnocení	nadprůměrná

Splnění cílů disertační práce	
Cíle disertační práce nejsou v práci sice explicitně uvedeny, ale jejich přehled je uveden v odst. 1. 3., kde jsou stručně shrnuty. Jejich splnění je pak velice skromně konstatováno v závěrečné kapitole práce. Vzhledem k tomu, že jde o vývoj speciální metody pro klasifikaci dat z dotazníků různých typů, se jedná o originální práci zasluhující mimořádné uznání. Bohužel citované výsledky není možné ověřit, protože programová řešení (kromě výpisu programů v textu práce) nejsou k dispozici a mohly být na vhodném paměťovém médiu přiloženy nebo případně na specifikované adrese, např. na některém serveru FD, dány oponentům a příp. dalším zájemcům k dispozici.	
hodnocení	nadprůměrné

Metody a postupy řešení	
V první kapitole práce je uvedena definice zvoleného tématu práce, současný stav jeho rozvoje, stručný přehled teoretických základů a struktura práce. Ve druhé kapitole jsou definovány základní pojmy, pravidla a modely, které jsou v práci dále používány, ve třetí a čtvrté kapitole jsou pak jednotlivé modely stručně popsány. V páté kapitole pak disertantka podrobně popisuje konstrukci lokálního modelu pro klasifikaci dat, klasifikaci dat naivním Bayesovým pravidlem a specifikaci speciálního algoritmu pro	

odhad parametrů a jejich klasifikaci marginálními datovými směsicemi (mixturami). V tomto smyslu se jedná o velmi významný výsledek. Přílohy práce, zejména přílohy B a D, které jsou podle mého názoru hodnotnější než základní text vlastní práce, pak vyčerpávajícím způsobem doplňují, resp. definují, odvozenou metodu, její význam pro praktické využití a rovněž hodnotí prezentované výsledky.

hodnocení

vynikající

Výsledky disertace – konkrétní přínosy disertanta

Výsledkem disertace je nově odvozená metoda založená na použití marginálních datových směsic a její experimentální ověření na simulovaných i reálných souborech dat. Jedná se o velmi významné výsledky, které budou moci být v budoucnu dále rozvíjeny. Z práce dále vyplývá, že disertantka prostudovala velké množství publikací vztahujících se k řešenému problému, řešenou problematiku a zejména dosažené výsledky v práci velmi pěkně vysvětluje (práce je dobře srozumitelná). Výsledky disertace jsou pak velmi hodnotné a přínos disertantky pro další rozvoj užitých metod značný.

hodnocení

vynikající

Význam pro praxi a pro rozvoj vědního oboru

Vývoj každé podobné metody a v práci uvedené výsledky mají pro praxi i pro další rozvoj vědního oboru velký význam, neboť řada dalších disertantů, případně i odborníků z praxe, bude moci na ně bez problémů navazovat – z tohoto pohledu jsou výsledky práce velmi významné.

hodnocení

nadprůměrný

Formální úprava disertační práce a její jazyková úroveň

Formální úprava disertační práce je velmi dobrá, v textu jsem našel jen dvě nepodstatné chybičky, které nijak nesnižují hodnotu práce, a svědčí o mimořádné pečlivosti autorky. Rovněž jazykovou úroveň práce nelze nijak kritizovat, anglický text vlastní práce i všech příloh je výstižný, srozumitelný a velmi dobře se čte, což pro mnoho podobných prací nebývá splněno.

hodnocení

nadprůměrná

Připomínky a dotazy:

Práce je zpracována pečlivě, s hlubokou znalostí věci a ilustrována vhodně zvolenými příklady, takže k textu nemám žádné připomínky. K práci mám jen několik drobných dotazů:

- str. 14 dole, bod 2b – the simulation is an imitation of the real situation (e.g. 2 or 3 modes) – jaké módy uvažujete ?
- str. 21 – jaký význam má označení T v T-junction ? Jde o časovou konstantu ? Na následující stránce užíváte T jako hodnotu času pro dosažení posledního vzorku.
- str. 27 – užíváte pojem pointer – označujete jím index nebo adresu komponenty ?
- str. 39, poslední řádka – components describe the clusters – vybrané nebo všechny ? A jak ?
- str. 53 dole – volíte $x_1 \in \{1, \dots, 4\}$ a $x_2 \in \{1, \dots, 6\}$ – na základě čeho? Na str. 56 pak volíte hodnoty 3 a 4.
- str. 58 – co označujete symboly xL a yL ?
- str. 60 dole – jak jste zvolila hodnoty parametrů komponent $C(i)p_j$?
- str. 72 – tab. 6.9 – čím jsou způsobeny rozdíly v hodnotách přesnosti klasifikace u aut, když v ostatních sloupcích jsou hodnoty zcela srovnatelné ?
- str. 92 – uvádíte „selected data are expertly assigned...“ – kdo byl expert ?
- str. 97 – uvádíte „The functions are necessary for the operation of the code after execution in SciLab, but they are not essential for the core of the thesis“ – proč ?

III. Závěrečné zhodnocení

Závěrečné zhodnocení disertace:

Předložená disertační práce má mimořádnou hodnotu, a to jak z hlediska návrhu nové metody, tak z hlediska její prezentace a ověření na vybraných příkladech. Rovněž formální provedení práce je v podstatě dokonalé, text je zcela srozumitelný, algoritmy jsou vhodným způsobem prezentovány. Nedostatkem je absence paměťového média se souborem zpracovaných programů, aby bylo možno funkci jednotlivých modulů vyzkoušet, resp. ověřit, a to případně i na jiných datech. Celkově lze práci hodnotit jako nadprůměrnou a bude velmi dobrým zdrojem pro řadu dalších doktorandů, kteří na ni, doufám, budou v budoucnu navazovat.

Udělení titulu Ph.D. **jednoznačně doporučuji.**

V Plzni dne 8. 9. 2023

A handwritten signature in blue ink, appearing to read 'Knapová', written over a horizontal dotted line.

podpis oponenta