



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:	doc. Ing. Jaroslav Sklenář, CSc.
Pracoviště oponenta práce:	Ústav telekomunikací, Fakulta elektrotechniky a komunikačních technologií, VUT Brno

II. Hodnocení jednotlivých kritérií

Aktuálnost tématu disertační práce	
Téma práce je velmi aktuální jak v kontextu neustále se zvyšujících objemů analyzovaných dat, tak také vzhledem k potřebě analyzovat diskrétní a často nominální, ordinální a multimodální data z různých dotazníků a podobných zdrojů.	
hodnocení	vynikající

Splnění cílů disertační práce	
Cíle práce byly splněny. Autorka popisuje novou metodu analýzy diskrétních multimodálních dat kdy za předpokladu nezávislosti jednotlivých proměnných a jejich rozdělením do shluků (clusters) lze výrazně urychlit analýzu která při použití běžných kategorických tabulek vzhledem k rychlému nárůstu počtu kombinací (curse of dimensionality) neumožňuje efektivní klasifikaci a predikci na základě dat z rozsáhlých souborů. Metoda je porovnána s několika alternativními klasifikačními metodami s velmi nadějnými výsledky.	
hodnocení	vynikající

Metody a postupy řešení	
Autorka nejdříve ve stručném úvodu definuje základní pojmy a vztahy dále použité při konstrukci klasifikační metody. Dále popisuje jednotlivé formy kategorických modelů. Následuje popis směsi (mixture) rozdělení, kde se autorka ne zcela přesvědčivě omezuje pouze na směsi binomických rozdělení. Ty jsou použity při konstrukci nezávislých shluků pro jednotlivé popisné proměnné které jsou základem lokálních modelů. Ty jsou dále použitím Bayesovy věty využity při konstrukci výsledného predikčního modelu. Zde oceňuji zařazení vhodně zvolených	



ilustrativních příkladů které čtenáři značně usnadňují pochopení principů. Metoda je dále testována pomocí simulovaných dat a tří typů dat získaných z reálných situací. Konkrétně jde o následky dopravních nehod, zájem o vozidla a predikci opakování po diagnóze rakoviny prsu. U reálných dat je metoda porovnána s šesti známými klasifikačními metodami.

hodnocení

vynikající

Výsledky disertace – konkrétní přínosy disertanta

Hlavním konkrétním přínosem práce je vytvoření nové klasifikační a predikční metody která může značně urychlit analýzu rozsáhlých souborů. Zde bych považoval za vhodné porovnat rychlosti jednotlivých klasifikačních metod a dále otestovat novou metodu na rozsáhlých souborech dat.

hodnocení

nadprůměrný

Význam pro praxi a pro rozvoj vědního oboru

Pro praxi jsou výsledky využitelné přímo. Bude však nutné metodu dále testovat.

hodnocení

nadprůměrný

Formální úprava disertační práce a její jazyková úroveň

Po formální stránce je práce vypracována velmi pečlivě, počet nedokonalostí je zanedbatelný. Rovněž angličtina je velmi dobrá bez chyb běžně se vyskytujících v podobných pracích. Bohužel jediným větším formálním nedostatkem je použití desetinných čárek.

hodnocení

nadprůměrný

Připomínky:

Během prezentace doporučuji podrobnější popis algoritmu jak detekovat shluky (zejména ty které se překrývají) v popisných datech.

III. Závěrečné zhodnocení**Závěrečné zhodnocení disertace:**

Cíle práce byly splněny, nedostatky nejsou závažné, výsledky jsou použitelné v pedagogické práci i dalším rozvoji vědního oboru.



Udělení titulu Ph.D. doporučuji.

V Brně dne 21.8.2023

.....
podpis oponenta