



POSUDEK MAGISTERSKÉ DIPLOMOVÉ PRÁCE

Posudek oponenta

Autor práce: Ing. Táňa Lančová
Název práce: Big data v sociologii
Autor posudku: Jiří Vinopal
Návrh klasifikace práce: velmi dobře

Základní charakteristika práce.

Jedná se o přehledovou práci. Jejím cílem je představit v sociologii doposud nepříliš frekventovanou oblast tzv. Big data.

KONCEPČNÍ STRÁNKA PRÁCE

Má práce jasně formulovaný záměr, cíl, výzkumnou otázku a odpovídají jim závěry?

Autorka jasně deklaruje přehledový charakter práce a ten způsobem zpracování naplňuje.

Má práce adekvátně stanovenou a jasně vyloženou metodu a postup řešení?

Postup řešení odpovídá záměru představit přehledovou stať a autorka systematicky postupuje v jednotlivých krocích od vzniku sféry Big data a její reflexi v sociologii, přes zdroje dat a metody analýzy až k problémům, se kterými se v případě sociologického využití BD musíme vyrovnávat.

Jsou jednotlivé části práce (zejm. teoretická a empirická) vyvážené a vhodně propojené, vytváří text konzistentní celek?

Práce působí jako celek představující specifickou výzkumně-metodologickou oblast, text nikde výrazněji nevybočuje z linie výkladu a drží se koncepce deklarované v úvodu.

Jsou argumentace a text práce a vystavěny logicky, jasně, srozumitelně? Má práce přehlednou strukturu?

Argumentace je vystavěna jasně a logicky, text má přehlednou strukturu a je přehledný.

Hodnocení kvality teoretické části práce (relevance k tématu a záměru práce, šíře a hloubka, vhléd autora do problematiky, samostatnost, originalita zpracování,...).

Z textu je velmi patrné, že autorka sféru Big dat důkladně prozkoumala jak z praktického hlediska možných typů a zdrojů konkrétních dat, tak z metodologického hlediska možných metod a způsobů jejich zpracování jak technického, tak statistického. V adekvátním rozsahu a přehledně provádí také reflexi dosavadního využití big dat v sociologii, resp. o úroveň výše dokonce metareflexi přijetí big dat v sociologii a jejich místo v metodologickém portfoliu.

Bohužel právě zde spatřuji největší slabinu práce, resp. přesněji řečeno její neúplnost. Práce jako taková má totiž především, řekněme, technickou povahu. Dobře ukazuje zdroje big dat, způsoby jejich zpracování i metody analýzy při použití v sociologickém výzkumu... Ovšem nepředstavuje možné cíle nebo výzkumné otázky, které by sociologie mohla jejich využitím řešit; neukazuje, v čem vlastně spočívá přidaná hodnota jejich využití z hlediska zvýšení poznání společnosti, které otázky, oblasti, problémy by se jimi daly řešit. Takto se celá věc jeví, že autorka (po mém osobním soudu oprávněně) pléduje za používání big dat v sociologii, ale jaksi vlastně sama neví, za jakým účelem by je tedy sociologie měla využívat, jaké konkrétní výzkumné otázky by se jimi mohla pokusit řešit. Nemyslím tím, že by autorka měla sama nějakou takovou analýzu v DP předvést, nemám nic proti přehledovému charakteru práce. Ale očekával bych (vedle všech stávajících) také informaci o tom, na jaký druh otázek, které se v sociologii objevují (anebo se třeba právě díky existenci BD teprve mohou nově objevit), může být analýza big dat prospěšná. Takový náčrt by mohl jít přes tematické oblasti (např. že jsou použitelná ve výzkumu sociálních otázek, politické sféry, kriminality a bezpečnosti, ekonomických otázek, hodnot, interakce a komunikace, důvěry atd.), anebo například přes úroveň složitosti: mohou big data pomoci s otázkami a problémy praktického rázu (aplikovaný výzkum), mohou objasnit jevy na úrovni teorií středního dosahu? Nebo mohou dokonce aspirovat na zodpovězení velkých otázek obecné sociologické teorie? Konečně, i kdyby ne takto, autorka mohla reálný potenciál alespoň ilustrovat například tím, že by si podle nějakého klíče vybrala několik konkrétních projektů sociologického výzkumu (např. těch aktuálně řešených v SOÚ AV ČR, který zmiňuje, nebo aktuálně financovaných GAČR) a představit, jaký přínos by pro jejich řešení mohlo využití BD představovat.

Absence představení praktických příkladů využití BD v sociologii mne mrzí o to více, že si v práci autorka cestu pro takovouto kapitolu připravuje: celkové rámování BD jako metodologické oblasti a inovace, která narušuje a zároveň nabízí originální řešení tradičního rozdělování kvantitativního a kvalitativního výzkumu, je pro diplomovou práci v daném oboru jednak výborně zvolené, jednak je to z mého pohledu jasná šipka pro pokračování právě ve směru ukázat, co konkrétně tedy takováto nová „metodologie“ může pomoci „klasické“ sociologii řešit.

Vzhledem k mladosti celé subdisciplíny BD a jejímu jen velmi pozvolnému etablování v sociologii by autorce v této době ještě nikdo jistě nevyčítal arbitrárnost, neúplnost nebo třeba i přehnaný optimismus ohledně představ o možnosti řešení konkrétních problémů. Naopak by čtenář v takovém případě musel ocenit skutečný velký vlastní přínos autorky a nezůstával by mu jako hlavní dojem z práce nakonec přece jen poněkud suchý klasifikační přehled typů, zdrojů a technických řešení.

Hodnocení kvality empirické části práce (adekvátnost použitých metod a postupů vzhledem k cíli práce, správnost aplikace metod a postupů, prezentace a interpretace výsledků,...).

Hodnocení empirické části není vzhledem k záměrně přehledovému charakteru textu relevantní. Z hlediska přehledové stati je postup výkladu a zpracování zcela v pořádku.

Hodnocení diskuse a závěrů práce (soulad interpretací a závěrů s výsledky empirických analýz, provedení diskuse závěrů, snaha o vysvětlení závěrů a jejich argumentace, zodpovězení výzkumných otázek, naplnění cílů práce).

Závěry práce odpovídají jejímu zaměření, autorka v nich de facto shrnuje potenciál big dat. I zde (anebo přinejmenším zde - v diskusi) mi však opět chybí nástin konkrétních výzkumných otázek, systematictější přehled příkladů možností konkrétního reálného využití BD v sociologii.

FORMÁLNÍ ASPEKTY PRÁCE

Hodnocení naplnění kritérií odborného textu (rozsah práce (BP 72-126 tis. znaků, DP 108-162 tis. znaků) citační normy, odkazy, popis datových zdrojů, jasnost odlišení myšlenek autora od převzatých, seznam literatury atd.)

Kritéria odborného textu jsou naplněna bez větších problémů.

Hodnocení vhodnosti a úplnosti využití informačních zdrojů (literární prameny, databáze, zahraniční literatura, datové zdroje,...)

Autorka používá vhodné informační zdroje.

Hodnocení jazykové, stylistické a grafické úrovně práce (grafická úprava, členění textu, označení příloh, zpracování tabulek, grafů, schémat atp.)

Jazyková a stylistická stránka práce vykazuje větší problémy. V textu jsou časté překlepy, objevují se nedokončené věty, text je místy příliš zkratkovitý, heslovitý.

Na s. 16 chybí poznámka pod čarou č. 11.

Na s. 23 jsou použity zkratky KVN a KVL, které nebyly nikde předtím definovány.

Celou kapitolu 6.4 by bylo dobré obohatit o příklady konkrétních typů úloh nebo řešených otázek; v této podobě jsou popisy metod, postupů a pojmů příliš obecné a málo pochopitelné pro člověka neznalého, pro znalého pak příliš stručné než aby mohly sloužit jako návody.

Z uvedených velmi stručných specifikací mi není úplně jasná povaha tří základních nejistot, „...se kterými se lze při analýze obecně setkat...“ dle Burrella (2016), které autorka uvádí na s. 55.

Na s. 57 se objevuje náznak toho, co kritizují výše – uvedení konkrétního příkladu využití (zde vysvětlované metody NN). Ovšem v tomto konkrétním příkladu mi chybí vysvětlení střední vrstvy neuronů..., o co konkrétně zde jde?

Co znamená „využití už dále nepodporovaného kódu („legacy code“) nebo vznik náhodných vzorců během simulace („weak emergence“),...“ na s. 67. Z uvedeného vysvětlení si stále nejsem úplně jistý porozuměním těmto problémům.

Další poznámky

Pobavila mne formulace „Po proběhnutí sítě volič zkolabuje do jednoho ze stavů volič strany xy,...“. ☺

Celkové hodnocení práce

Práce jako celek je konzistentní a dobře představuje aktuální stav využívání sféry Big dat v sociologii. Zůstává podle mne však za svým vlastním cílem předvést užitečnost této sféry, neboť čtenáři představuje jen velmi málo konkrétních námětů výzkumných otázek, cílů nebo problémů, při jejichž řešení by bylo využití BD užitečné. A to je pro nalákání zájemců podle mne nejdůležitější... Říct, „hele, tady je sklad plný kladívek, ty se skládají z dřevěného nebo umělého topůrka, mají kulatou nebo plochou hlavu, nějaké ti půjčí hlídač, když si to s ním domluvíš, a pak už ho použiješ prostě tak, že ho za topůrko chytneš do pravé ruky...“, to podle mne zkrátka nestačí. Abych pro kladívko do skladu vážil cestu, musí mi někdo v první řadě říct, že je to skvělá věc na zatlučení hřebíku! Anebo na roztřískání okna sousedovi;-) A protože v práci právě informaci tohoto druhu postrádám, ačkoli vše ostatní je zpracováno velmi dobře, nenavrhuji hodnocení výborně, nýbrž „pouze“ velmi dobře.

Otázky a náměty k obhajobě

U obhajoby bych ocenil zejména doplnění toho, co popisují, že mi v práci chybí – nějaká systematická reflexe konkrétních cílů/otázek/oblastí/úrovní/problémů, pro jejichž řešení v sociologii může být zapojení BD prospěšných.

Podobně bych ocenil také pokus o představení konkrétního výzkumného problému/projektu (mám na mysli návrh/vymyšlení vlastního, ne nutně vyhledání nějakého reálně běžícího), na kterém by se dobře vyjevila překlenovací povaha BD z hlediska kvantitativního a kvalitativního výzkumu.

Datum: 22. 8. 2018

Podpis: