

Posudek oponenta závěrečné práce

České vysoké učení technické v Praze

Fakulta informačních technologií

Student: Bc. Evgeniya Nenenko
Oponent práce: Mgr. Tomáš Lancinger
Název práce: Analysis of bank data
Obor: Znalostní inženýrství

Datum vytvoření: 6. 6. 2017

Hodnotící kritérium:	Způsob hodnocení - následující škálou 1 až 5:
1. Náročnost a další komentář k zadání	1=mimořádně náročné zadání, 2=náročnější zadání, 3=průměrně náročné zadání, 4=lehčí, ale ještě dostatečně náročné zadání, 5=nedostatečně náročné zadání
Popis kritéria: Podrobněji charakterizujte diplomovou (bakalářskou) práci a její případné návaznosti na předchozí nebo běžící projekty. Dále posuďte, čím je zadání této ZP náročné. (U obtížnější ZP lze dále tolerovat některé nedostatky, které by u ZP standardní obtížnosti tolerovány nebyly; a naopak u jednoduché ZP mohou být zjištěné nedostatky hodnoceny přísněji.)	
Komentář: Diplomová práce adresuje téma z komerčního prostředí, ve kterém je potřeba pracovat s řadou reálných omezení. Zadání je formulováno obecněji a koresponduje se skutečným zadáním, užívaným v bance. Práce je exaktní, závěry jsou relativně dobře vyargumentované a srozumitelné jako po technické, tak i obecné stránce. Diplomantka prokázala schopnost kompletního zpracování analytického problému - od uchopení problematiky až po dodávku provozovatelného zdrojového kódu a komunikaci výsledku. Mnohé její postřehy a věcné náměty kvalitativně posunuly původně předpokládaný rozsah práce, což hodnotím velice pozitivně.	
Hodnotící kritérium:	Způsob hodnocení - následující škálou 1 až 4:
2. Splnění zadání	1=zadání splněno, 2=zadání splněno s menšími výhradami, 3=zadání splněno s většími výhradami, 4=zadání nesplněno
Popis kritéria: Posuďte, zda předložená ZP splňuje zadání. V komentáři uveďte body zadání, které nebyly zcela splněny, případně rozšíření ZP oproti původnímu zadání. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.	
Komentář: Zadání shledávám jako splněné. Vyslovené hypotézy byly prověřeny korektním postupem a výsledek splňuje očekávání jak po věcné, tak i po formální stránce.	
Hodnotící kritérium:	Způsob hodnocení - následující škálou 1 až 4:
3. Rozsah písemné zprávy	1=splňuje požadavky, 2=splňuje požadavky s menšími výhradami, 3=splňuje požadavky s většími výhradami, 4=nesplňuje požadavky
Popis kritéria: Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části.	
Komentář: Práce splňuje požadovaná kritéria ohledně rozsahu, struktury a relevance.	
Hodnotící kritérium:	Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):
4. Věcná a logická úroveň práce	90 (A)
Popis kritéria: Posuďte, zda předložená ZP je po věcné stránce v pořádku, případně vyskytují-li se v práci věcné chyby nebo nepřesnosti. Zhodnoťte dále logickou strukturu ZP, návaznosti jednotlivých kapitol a pochopitelnost textu pro čtenáře.	

Komentář:

Práce představuje postupy pro detekci vzorů v transakčních datech (časových řadách) a ověřuje jejich funkčnost na reálných datech.

Po věcné stránce tedy dosáhla vytyčeného cíle, kterým bylo zejména ověření vyslovených hypotéz, zda lze (a do jaké míry) zjistit společné vzory transakčního chování firem a odvodit z nich společné charakteristiky, které mohou vést na nová zjištění a znalosti, zlepšující zaměření obchodních aktivit banky.

Strukturování práce je poměrně logické a intuitivní.

Snad jedině, co bych vnímal jako možný prostor pro zlepšení bylo detailnější rozvedení potenciálu pro detekci segmentových meta-patternů, které by umožnily de facto do jisté míry zrekonstruovat expertní znalost seniorních obchodníků. Zmíněný aspekt nicméně nevnímám jako výhradu, spíše jen jako námět na případné další směřování a rozvoj - striktně vzato ani nebyly ve výchozích hypotézách.

Hodnotící kritérium:

Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):

5. Formální úroveň práce

90 (A)

Popis kritéria:

Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 14/2015, článek 3.

Komentář:

Práce je psána v angličtině a je veskrze dobře srozumitelná, byť v některých pasážích by se nabízely plynulejší anglické obraty, usnadňující čtení a pochopení textu.

Z pohledu jazykové stránky a formulace představené myšlenky ji shledávám jako plně vyhovující.

Hodnotící kritérium:

Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):

6. Práce se zdroji

95 (A)

Popis kritéria:

Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení ZP. Charakterizujte výběr studijních pramenů. Posuďte, zda student využil všechny relevantní zdroje nebo zda se pokoušel řešit již vyřešené problémy. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.

Komentář:

Práce obsahuje 18 bibliografických referencí, které jsou v textu používány důsledně. Citované a převzaté pasáže a ilustrace jsou řádně označené.

Hodnotící kritérium:

Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):

7. Hodnocení výsledků, publikační výstupy a ocenění

95 (A)

Popis kritéria:

Vyjádřete se k úrovni dosažených hlavních výsledků ZP, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, apod. Případně také zhodnoťte, zda software nebo zdrojové texty, které nevytvořil sám student, byly v ZP použity v souladu s licenčními podmínkami a autorským právem. Popište případnou publikační činnost a získaná ocenění související s řešením této ZP.

Komentář:

Diplomantka dokázala vybrat, implementovat a rozšířit algoritmus, který splňoval zadání pro použití v analýze pro podporu obchodních procesů.

Realizovaný program bude nasazen do Data Science Laboratoře. Sice momentálně nepodporuje masivně paralelní běh, ale tato okolnost byla zvažena (a v diplomové práci je tato problematika popsána a vyjasněna) a na doporučení Data Science zanedbána, protože algoritmus bude mít k dispozici dostatečný strojový výkon pro plnohodnotný běh i bez optimalizací.

Hodnotící kritérium:

Způsob hodnocení - nehodnotí se

8. Komentář o využitelnosti výsledků

Popis kritéria:

Uvedte, zda hlavní výsledky ZP rozšiřují již publikované známé výsledky a/nebo přinášející zcela nové poznatky. Uvedte možnosti využití výsledků ZP v praxi.

Komentář:

Výsledky diplomové práce jsou plně využitelné pro analytické úlohy Data Science v Komerční bance. Realizovaný algoritmus dokáže detekovat významné vzory chování firem, což jsme si ověřili na vybraném reprezentativním vzorku.

Z pohledu praktičnosti bude muset být stávající systém doplněn o interpretační modul, který ze zjištěné databáze vzorů chování sestaví uchopitelnější (méně technické) výsledky. Případně bude aplikace poskytovat vstupní data do dalšího analytického systému.

Hodnotící kritérium:

Způsob hodnocení - nehodnotí se

9. Otázky k obhajobě

Popis kritéria:

Uvedte případné dotazy, které by měl student zodpovědět při obhajobě ZP před komisí (body oddělte odrážkami).

Otázky:

Mohla byste rozvést návrh na přidání obchodního významu zjištěným vzorům, který zmiňujete v závěru v sekci Future work?

Hodnotící kritérium:

Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):

10. Celkové hodnocení

95 (A)

Popis kritéria:

Shrňte stránky ZP studenta, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení **nemusí** být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích 1 až 9.

Text hodnocení:

Diplomantka prokázala schopnost systematického uchopení zadané problematiky - provedla rozbor zkoumaných hypotéz, provedla rešerši nabízených řešení, kvalifikovaně zhodnotila a vybrala nejlepší přístup a následně zvolený algoritmus implementovala a zpracovala s ním poskytnutá anonymizovaná data. Navíc byla schopna do algoritmu vnést nová rozšíření pro lepší adaptaci na specifický charakter zadaní. Výstup z její diplomové práce bude zařazen do analytického portfolia oddělení Data Science Komerční banky a bude na něm prováděn další rozvoj pro postupnou produkcionizaci.

Proto hodnotím celou práci jako velmi dobře zvládnutou a konstatuji, že z mého pohledu byl zadaný cíl diplomové práce splněn.

Podpis oponenta práce: