

POSUDEK OPONENTA BAKALÁŘSKÉ PRÁCE

Jméno studenta: Petr Hollman

Název bakalářské práce: Aplikace dolování dat pro podporu rozhodování v moderních podnicích

Jméno vedoucího bakalářské práce: Ing. Athanasios Podaras, Ph.D.

Jméno oponenta bakalářské práce: Ing. Šimon Kolář, MiM, FRM

Oponent: - **název firmy:** KPMG Česká republika, s.r.o.

- **pracovní zařazení:** Manager, Risk Consulting **kontakt:** skolar@kpmg.cz

	Výborně	Velmi dobře	Dobře	Neprospěl
I. Hodnocení zpracování tématu studentem:				
Splnění cíle práce		X		
Volba metod a jejich aplikace při zpracování tématu	X			
Hloubka provedené analýzy		X		
II. Hodnocení struktury a obsahu práce:				
Přehlednost a logická stavba (struktura) práce		X		
Aktuálnost a vhodnost použitých pramenů	X			
Schopnost studenta zpracovat získané podklady	X			
Přiměřenost a srozumitelnost závěrů práce	X			
Formulování vlastních názorů studenta		X		
III. Hodnocení formy a stylu práce:				
Formální úprava práce (text, tabulky, grafy)			X	
Stylistická úroveň práce			X	
Práce s českou literaturou včetně odkazů a citací	X			
Práce se zahraniční literaturou včetně odkazů a citací		X		

Slovní hodnocení a připomínky k bakalářské práci v rozsahu minimálně deseti řádků uveďte, prosím, na druhou stranu posudku.

Otázky k obhajobě bakalářské práce:

1. Jaké další existující problémy v sektoru bankovníctví lze podle Vás řešit metodami Data Miningu a jaké metody a nástroje by se daly k jejich řešení využít?
2. Jaké existující aplikace byste doporučil malé až střední finanční instituci využít k řešení Data Miningových úloh?

Práci doporučuji - nedoporučuji* k obhajobě. (*nehodící se škrtněte)

Bakalářskou práci navrhuji klasifikovat stupněm: Velmi dobře

Datum: 20.5.2022


Podpis oponenta bakalářské práce

Předložená práce se věnuje problematice Big Data, primárně metodám zpracování tzv. velkých dat a jejich aplikaci na reálném příkladu, výpočtu FLI faktoru v rámci modelu očekávané úvěrové ztráty dle požadavků Mezinárodního standardu finančního výkaznictví IFRS 9. Téma je aktuální a prakticky relevantní, IFRS 9 vstoupil v platnost 1. 1. 2018 a tematika stanovení FLI faktoru je stále dostatečně živá a různorodá. Je rovněž dostatečně komplexní, vzhledem k šíři používaných přístupů a praktickým obtížím při stanovení tohoto faktoru, a je tedy vhodným tématem pro práci tohoto rozsahu. Zpracování tohoto tématu vyžaduje jak znalost základních metod analýzy dat, tak požadavků standardu IFRS 9, a v případě praktické části rovněž volbu vhodné metody a jí odpovídajících dat, jedná se tedy o poměrně obtížné, multidisciplinární téma.

Práce jako celek je poměrně rozsáhlá. V prvních částech se věnuje obsáhle problematice Big Data a Data Mining, včetně definice a historie těchto pojmů a práce s nimi, jejich základním vlastnostem, klasifikaci a popisu různých technik a používaných nástrojů. Následně obrací svou pozornost na stručný úvod do problematiky výpočtu očekávané úvěrové ztráty dle IFRS 9. Zde autor stručně popisuje veškeré podstatné základní pojmy a jejich vztah, kapitola však obsahuje několik zjednodušení a nepřesností – například prezentovaný model založený na rizikových parametrech je pouze jedním z možných přístupů připouštěných standardem, což není nikde zmíněno. Autorem popsany způsob aplikace FLI coby úpravy PD je rovněž pouze jednou z variant zohlednění tohoto požadavku IFRS 9. Pro účely práce je však prezentované shrnutí dostatečné. V praktické části pak autor prezentuje jeden z přístupů stanovení FLI metodou vícenásobné lineární regrese, jež odhaduje vliv vybraných makroekonomických veličin na parametr PD a následné stanovení FLI koeficientu upravujícího parametr PD pomocí vážených hodnot vývoje modelovaného PD v několika scénářích, stanovených pomocí vybraných percentilů analyzovaných veličin. Prezentovaný přístup patří mezi relevantní, v praxi používané přístupy stanovení FLI faktoru. Tento přístup je současně založen na relevantních datech a jejich statistické analýze a dostatečně jednoduše prakticky uchopitelný na to, aby mohl být inspirací pro menší finanční instituce. Zvolený příklad rovněž umožňuje v míře omezené, avšak pro účely práce dostatečné, předvést v první části práce prezentované postupy práce s daty. Autor zde prezentuje relevantní statistické testy a jejich interpretaci, stejně tak interpretaci výsledků modelu. Bylo by však vhodné explicitněji provázat postupy prezentované v první části s praktickou částí. Zásadnější je však chybná interpretace některých statistických testů – prezentovaný výsledek Bresch-Paganova testu při zvolené hladině významnosti svědčí o nesplnění předpokladu homoskedasticity, ačkoliv autor tvrdí opak. U Kolmogorova-Smirnova testu je chybně uvedeno, že jeho výsledek vyvrátil hypotézu normálního rozdělení vysvětlujících proměnných, ačkoliv právě normalita proměnných je předpokladem lineární regrese. A konečně v kapitole Interpretace koeficientu FLI prezentované percentily pro jednotlivé scénáře v případě proměnných inflace a nezaměstnanost neodpovídají vztahům těchto proměnných s vysvětlovanou proměnou jež vyplývají z regresního modelu, aniž by byl tento fakt uspokojivě vysvětlen. Na druhou stranu, robustní statistická analýza přesahuje zvolené téma i obor autora.

Celkovou úroveň zpracování hodnotím jako ucházející. K silným stránkám práce patří její praktická relevantnost a komplexní uchopení relativně složitého odborného tématu. K nedostatkům, mimo výše uvedených nepřesností, patří zejména její formální podoba. Práce obsahuje řadu formátových a jazykových nepřesností, několik odstavců je například patrně nedopatřením uvedeno dvakrát, některé pojmy jsou zaměňovány, některé použité zkratky nejsou uvedeny v seznamu zkratek atd. Bylo by rovněž vhodné použít při rešerši více odborné literatury na úkor internetových zdrojů. Přes uvedené nedostatky práci doporučuji k obhajobě, zejména s ohledem na komplexnost tématu a jeho velmi praktické uchopení, jež svědčí o dobré schopnosti autora provázat teoretické znalosti se skutečnými problémy podnikové praxe.