

## POSUDOK ZÁVEREČNEJ PRÁCE

**Téma:** Návrh modelu na bázi Soft Case-based Reasoning

**Typ záverečnej práce:** Doktorandská záverečná práca

**Študijný program:** Systémové inžinierstvo a informatika

**Študijný obor:** Informatika vo verejnej správe

**Autor:** Ing. Filip Mezera

**Oponent:** doc. RNDr. Michal Munk, PhD.

Práca sa zaoberá úlohou klasifikácie, kde autor na riešenie tohto problému kombinuje prípadové usudzovanie s ďalšími analytickými metódami (metódy strojového učenia, štatistické metódy, fuzzy/rough množiny).

Práca je dobre čitateľná, pomerne náročná problematika je zrozumiteľne popísaná. Kladne hodnotím, že práca má aplikačné domény, výsledky je možné použiť na riešenie konkrétnych problémov v doméne financií, či kvality ovzdušia. Oceňujem dotiahnutie teoretickej práce až do aplikačných výstupov. Autor preukázal, že pri riešení problémov/úloh vie vhodne kombinovať rôzne nástroje za účelom predspracovania a analýzy dát, čo je typická prax v oblasti získavania znalostí.

V kapitole 1 je podrobne rozpísaný cieľ práce ako aj úlohy (čiastkové ciele) potrebné na dosiahnutie stanoveného cieľa. Úlohy (vyplývajúce z cieľa práce) sú ďalej podrobne rozpísané a čiastočne nahrádzajú absentujúcu metodiku práce.

Kapitola 2 popisuje použité metódy, konkrétne v prvej časti prípadové usudzovanie- modely a fázy a v druhej fuzzy/rough množiny a neurónové siete. Analytické metódy (fuzzy/rough množiny a neurónové siete) sú spočiatku popísané okrajovo v súvislosti s prípadovým usudzovaním, v závere popisu jednotlivých metód však autor udáva ich využitie v prípadovom usudzovaní.

Kapitoly 3, 4 predstavujú nosnú časť práce, kde boli riešené úlohy klasifikácie v rôznych doménach s cieľom porovnať skúmané metódy klasifikácie. Výsledky popísané v kapitolách 3

a 4 boli publikované v časopise s nenulovým SJR a v sériách vydávaných vydavateľstvom Springer.

V práci absentuje metodika, ktorá na jednom mieste sprostredkováva postup od získania dát, cez predspracovanie dát, modelovanie, krížovú validáciu až k vyhodnoteniu a porovnaniu výsledkov získaných rôznymi prístupmi k riešeniu úlohy klasifikácie v oblasti kvality ovzdušia a financií.

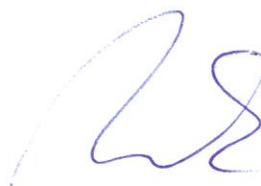
Práca je precízne spracovaná, obsahuje iba menej závažné formálne nedostatky: matematické výrazy by mali byť tiež súčasťou viet; nejednotná forma popisku obrázkov/tabuliek (veľké/malé písmeno); zamieňanie skratiek RST/RTS pre Rough Sets Theory, podobne RFA/FSA pre Rough-Fuzzy Approach.

Ak porovnáme parametrické a neparametrické postupy, tak za predpokladu, že sú splnené predpoklady použitia parametrickej metódy je táto metóda robustnejšia, resp. ak sa k rovnakým výsledkom dopracujeme parametrickými ako aj neparametrickými postupmi môžeme ich považovať za robustné. Avšak na robustnosť sa dá nazerať rôzne, objasnite ako vo vašom prípade chápete robustnosť metódy, resp. výsledkov?

Celkovo prácu hodnotím ako veľmi dobrú, navyše s aplikačným potenciálom. O čom svedčia aj publikačné výstupy autora.

Práca spĺňa požiadavky kladené na tento typ záverečnej práce a prácu odporúčam v predloženej podobe obhajovať a po jej úspešnej obhajobe navrhujem, aby Ing. Filipovi Mezerovi bol udelený akademický titul Philosophiae Doctor (PhD.) v študijnom programe Systémové inžinierstvo a informatika, oboru Informatika vo verejnej správe.

V Nitre, 20.05. 2018



doc. RNDr. Michal Munk, PhD.

# Oponentský posudek dizertační práce

Název práce: ***Návrh modelu na bázi Soft Case-based Reasoning***  
Dizertand: *Ing. Filip Me z e r a*  
Univerzita Pardubice, Fakulta ekonomicko-správní  
Obor: Informatika ve veřejné správě  
Školitel: Doc. Ing. Jiří Křupka, PhD.,  
Univerzita Pardubice, Fakulta ekonomicko-správní  
Oponent: Prof. Dr. Ing. Miroslav Pokorný, VŠB -Technická univerzita Ostrava,  
Fakulta elektrotechniky a informatiky.

Předložená dizertační práce je rozčleněna do 6 pracovních kapitol vč. úvodu, cílů práce, závěru a standardních doplňků. Práce je napsána na 83 stranách. Kapitola 2 je věnována analýze současného stavu a možností řešení, kapitoly 3 a 4 tvoří tvůrčí část dizertační práce.

## *Dosažení stanovených cílů*

Cíle dizertační práce jsou explicitně deklarovány v Kapitole 1. Za izertabilní považují návrh problémově orientovaných klasifikačních modelů s využitím nekonvenčních metod soft-computingu s důrazem na využití neurčitosti informací, analýzu vlastností modelů a stanovení předpokladů pro jejich úspěšné následné použití. Posouzením průběhu řešení a jeho výsledků konstatují, že stanovené cíle byly splněny.

## *Postup řešení problému*

Téma práce je zpracováno s velkou šíří záběru. Struktura dílčích cílů je v rámci tématu dizertační práce kompaktní a dobře promyšlená. K řešení jsou využity konvenční i nekonvenční metody abstraktního modelování i posouzení kvality predikcí a klasifikací. Vybrané metody řešení tématu jsou adekvátní, forma a obsah jejich analýzy je dobrým základem pro zdůvodnění jejich použití při řešení tématu práce. Dizertační práce je charakteristická jak svojí formou prezentace, tak i svým obsahem. Z formálního hlediska je volen popis postupů řešení a návrhů modelů bez prezentace schémat jejich struktur a uvedení jejich parametrů. Schémata se omezují pouze na jejich ideové řešení, bližší informace jsou řešeny odkazy na literaturu. Tato forma prezentace použitých metod a nástrojů nedává představu o jejich konkrétní formalizaci doktorandem v rámci řešení a modely nelze z tohoto hlediska blíže posoudit. Autor uvádí analýzy a tabelované nebo grafické numerické výsledky testování a verifikací. Z tohoto hlediska má dizertační práce spíše charakter komentovaných výsledků dosavadních prací, které však nejsou jejími přílohami. Absence detailnějších popisů omezuje jak možnosti posouzení řešení oponentem, tak také využití práce v oblasti výuky.

Z hlediska obsahu práce je typické širší pojetí způsobu řešení tématu, spočívající ve směru výběru dvou typů klasifikačních modelů ze dvou různých oblastí, z nichž žádný nelze považovat za triviální. Doktorand v rámci výzkumu provádí analýzu vlastností různých modelovacích metod s cílem verifikace jejich kvality a robustnosti. Z tohoto hlediska je třeba také podtrhnout, že doktorand věnuje velkou pozornost vlastnostem použitých datových souborů.

Za velmi zajímavou považuji část, týkající se kombinace jednotlivých metod v hybridních modelech. I když je metodologie takového řešení známá z literatury, musel doktorand pro její praktickou realizaci v rámci konkrétních úloh prokázat značný stupeň kreativity.

Pro soft-computingový hybridní dynamický model byly vzaty v úvahu metody expertních fuzzy-logických systémů, genetických algoritmů, případového usuzování a umělých neuronových sítí.

Na základy zhodnocení klasifikačních a predikčních vlastností předchozích modelů navrhl doktorand dva původní hybridní modely – hybridní model dynamické klasifikace a hybridní model klasifikace klienta. Oba tyto modely jsou prezentovány opět pouze ideovými blokovými schématy. Pro kombinaci metod byly zvoleny přístupy případového usuzování a vícevrstvých neuronových sítí.

V rámci řešení modelu dynamické klasifikace je zajímavá analýza možností získání klasifikačních pravidel. Doktorand ověřil výsledky s využitím přístupů expertního odhadu a hrubých množin.

Metodologie návrhu struktur hybridních modelů jsou – podle mého názoru – kvalitním závěrem řešení tématu dizertační práce, mají význam pro praktické postupy v obdobných případech a jsou dobrým východiskem dalšího výzkumu.

#### *Význam pro rozvoj oboru a praxi*

Z hlediska rozvoje teorie oboru aplikované informatiky považuji za přínosné syntézy a analýzy klasifikačních modelů, založených na metodě případového usuzování s využitím moderních metod výpočtové inteligence. Práce je přínosem k porovnání vlastností modelů, využívajících ke klasifikaci na jedné straně generalizovaná pravidla, na druhé straně relevantní případy z historie. Významné je i rozpracování přístupu validace modelů s využitím takových metod, jako jsou rozhodovací stromy, umělé neuronové sítě, shlukování, logistická regrese nebo citlivostní analýza.

Z praktického hlediska jsou zajímavá zaměření aplikačních oblastí vybraných modelů, tedy predikce vývoje kvality ovzduší a řešení smogových situací a klasifikace klientů obchodních organizací z hlediska jejich dlouhodobého přínosu. Výsledky zlepšení kvality predikce a klasifikace oproti aktuálně užívaným modelům potvrzují okamžitý přínos řešení v jejich praktickém uplatnění. Hodnotím rovněž schopnosti doktoranda v oblasti efektivního využívání informatických metod a nástrojů.

#### *Formální úprava a jazyková úroveň*

Práce má velmi dobrou jazykovou úroveň, její grafická stránka je adekvátní obsahu. Text je prakticky bez pravopisných chyb.

#### *Publikační aktivita doktoranda*

Výsledky řešení tématu dizertační práce byly průběžně publikovány. Publikace jsou rovnoměrně rozděleny do celé doby tvůrčí části řešení. Publikace jsou skromnější, nicméně dizertabilní jádro práce bylo řádně zveřejněno. Ve struktuře publikací chybí příspěvky ve sbornících zahraničních konferencí.

#### *Otázky oponenta*

Otázky jsou soustředěny do části řešení, v níž jsou přiloženy konkrétní struktury a parametry modelu (model kvality ovzduší).

1. Příklad 7 – tvary funkcí příslušnosti přiřazují určitou hodnotu výstupní proměnné  $y$  jednoznačně k jazykové hodnotě LOW, přitom ale současně umožňují její nenulovou příslušnost i k jazykovým hodnotám VERY LOW a NORMAL. Tatož syntaxe je použita i v případě přiřazování vstupních hodnot v Příkladu 6.

2. 3840 pravidel šesti-atributového fuzzy modelu degraduje výstupní jazykovou proměnnou pouze do dvou jazykových hodnot GOOD-BED, definovaných na spojitém normovaném univerzu  $(0,1)$ . Jakým způsobem bylo podle defuzzifikované hodnoty ostře, tedy dvouhodnotově, rozhodnuto, zda stav ohrožení vyhlásit či ne?

#### *Závěrečné hodnocení a doporučení*

Předložená dizertační práce pana Ing. Filipa Mezery prokazuje jeho dobrou odbornou orientaci a schopnost řešit úlohy v problematice tvorby modelů komplexních soustav, jejich implementace a analýzy. Ing. Filip Mezera prokázal při řešení tématu dizertační práce teoretické znalosti, schopnost odborné práce i schopnost aplikace jejích výsledků. I když dokumentační stránka práce je oslabena, doporučuji ji k obhajobě. Bude-li obhajoba úspěšná, doporučuji, byl panu Ing. Filipovi Mezerovi, v souladu s §72 Zákona 111/98 Sb. o vysokých školách a příslušnými ustanoveními předpisu o doktorském studiu na Univerzitě Pardubice a její Fakultě ekonomicko-správní, udělen akademický titul

**„D o k t o r, Ph.D.“**

v akreditovaném oboru doktorského studia Informatika ve veřejné správě.

Ostrava, 10.5.2018



Prof. Dr. Ing. Miroslav P o k o r n ý

## Oponentní posudek disertační práce

### Název práce: Návrh modelu na bázi Soft Case-based Reasoning

Autor: Ing. Filip Mezera

Školitel: doc. Ing. Jiří Křupka, PhD.

Oponent: prof. RNDr. PhDr. Antonín Slabý, CSc., Fakulta informatiky a managementu, Univerzita Hradec Králové

### Struktura a obsah práce

Práce s rozsahem 99 stran je věnována problematice případového usuzování (Case-based Reasoning, dále jen CBR) s využitím metod výpočetní inteligence (Computational Intelligence,

Základní text práce je po obligátních seznamech obrázků, tabulek, a zkratk a Úvodu členěn na 8 číslovaných kapitol, z toho Kapitola 7 je seznamem použité literatury a Kapitola 8 obsahuje přílohy práce.

Stručná Kapitola 1 – Cíle práce formuluje cíle a velmi stručně se věnuje použitým metodám a důvodům jejich použití v modelech. Kapitola 2 - Popis soudobého stavu řešené problematiky vymezuje současné způsoby řešení a modely CRB a zejména se věnuje v práci použitým CRB modelům - Huntovu modelu CBR a Modelu 4R a ve druhé části ukazuje 3 fáze CRB a 3 hlavní nástroje použité k řešení problematiky (fuzzy množiny, neuronové sítě a rough množiny). Následují kapitoly věnované dvěma podstatným a z hlediska postupu a metod odlišným aplikacím a podrobnému popisu způsobu tvorby jejich modelu a jejich řešení.

Kapitola 3 se týká Klasifikačního modelu kvality ovzduší a celému postupu jeho vývoje a aplikaci použitých metod. Kapitola 4 je zaměřena na Klasifikační modely v nebankovní finanční instituci. Kapitola 5 - Naplnění cílů práce je věnována prezentaci výsledků. Následuje kapitola 6 - Závěr a obligátní Kapitola 7 - Literatura s přehledem použitých zdrojů a Kapitola 8 – přílohy.

### Cíl práce, výzkumné otázky, aktuálnost tématu, soulad s oborem studia

Cílem práce je navrhnout klasifikační modely s využitím metody Soft CBR (SCBR) a ověřit jejich schopnost pracovat s neurčitostí v datech. Ve dvou aplikacích je zkoumán v prvním případě statický klasifikační model, v druhém případě dynamický model. Jsou testovány různé přístupy v rámci jednotlivých fází CBR. Dosažené výsledky jsou porovnány s výstupy, které byly dosaženy pomocí dalších metod (například NNs, TDIDTs, RST). Hlavní cíl je rozdělen na následující subcíle:

- Stanovení vhodných metod omezujících počet vstupních atributů.
- Cílené využití metod definujících strukturu báze případů.
- Porovnání různých přístupů k získání nejpodobnějšího případu / případů a jejich vliv na výslednou klasifikaci
- Vytvoření dynamického modelu CBR
- Zjištění možnosti integrace výsledků jednotlivých klasifikačních metod včetně CBR, tedy vytvoření hybridního inteligentního systému.

Dílčí kroky jsou správně voleny a vedou k dosažení cíle. Téma a cíl práce jsou zajímavé a užitečné, náročné a velmi aktuální. Práce svojí tematikou patří do oboru doktorského studia Systémové inženýrství a informatika..

## **Použité vědecké metody v disertační práci**

Práce vychází z rešerše a analýzy literárních zdrojů o předmětné problematice. Jsou využity a zpracovány práce reálné datové zdroje. V práci je použita řada specifických metod - bez nároku na úplnost jmenujme aspoň Huntův model případového usuzování a k jeho optimalizaci torie rough množin a fuzzy množin využité v rámci 1 aplikace - modelování kvality ovzduší, model případového usuzování 4R byl použit a upraven využitím neuronových sítí a rough množin za použití různých metrik v rámci druhé aplikace - modelu klasifikace nových klientů.

Při hodnocení výsledků byly pro porovnání využity rozhodovací stromy a logistická regrese. V rámci práce je použita řada dalších speciálních metod.. Jsou zařazeny / zmíněny i speciální metody patřící do oper. výzkumu (hill climbing apod.), fuzzy přístupy, speciální metody kontrolní (shlukování, log regrese a pod.). Dále použil autor klasické a osvědčené obecné vědecké metody a přístupy- např. deskripce, klasifikace, kritické zhodnocení

Potenciál, který mají metody a jejich skupiny pro řešení v rámci konkrétní aplikace je v práci podrobně a správně rozebrán. Metody je možno v souhrnu považovat za adekvátní, správné, a ke splnění konkrétních cílů vedoucí, rozhodnutí o použití konkrétní metodiky je nutno považovat za netriviální. Pro výpočty je použito několik softwarových aplikací systém Matlab apod.

## **Splnění cílů práce**

Cíle práce a tedy i dílčí cíle k němu vedoucí byly splněny. K hlavnímu cíli práce a podobně ke všem dílčím krokům vedoucím k jeho dosažení přispěl autor věrohodným příspěvkem. Postupy a metody jsou vybrány adekvátně a aplikovány účelně, a metodický rámec je rámcově opakovatelně použitelný.

## **Přesnost práce, formální stránka práce**

Práce je napsána stručným, jasným a dostatečně přesným jazykem. Formální stránka práce i přesnost vyjadřování, úprava výstupů (vzorců, diagramů obrázků a tabulek) jsou na odpovídající úrovni. Práce má jasnou strukturu, proto je možno se v ní dobře orientovat.

## **Výsledky práce a poznatky a přínosy práce**

Práce přináší výsledky v oblasti teoretické, metodologické i praktické. Práce také jedinečným způsobem zúročuje autorovy pracovní zkušenosti a zájmy.

Za hlavní teoreticko-metodologické výsledky lze považovat tvůrčí syntézy provedené v rámci metodiky a řešení 2 aplikací zkoumaných v práci:

Huntova modelu případového usuzování a torie rough množin a fuzzy množin v případě první aplikace a modelu případového usuzování 4R a využití neuronových sítí a rough množin v případě druhé aplikace. Celé metodiky postupu v obou případech s řadou podstatných implementačních detailů mají samostatnou hodnotu.

Za přínos lze považovat i pěkné a srozumitelné a dosti komplexní a v některých směrech hluboké zpracování tématiky s použitím adekvátních metod.

Práce může mít i rozsáhlé potenciální využití praktické. Nabízí se (modifikované, tvůrčí) opakované využití metodických postupů uvedených v práci v případě deterministického i stochastického modelu.

## **Publikační činnost autora ve vztahu k práci**

Publikační činnost autora zahrnuje celkem 5 položek ve zdrojích adekvátních tématu práce.. Pozitivním rysem je navázání tematiky práce i publikačních aktivit na výzkum školitele a pracoviště.

## **Otázky do diskuse obhajobě**

Práce ukázala zároveň, že problematika CBR je náročná, široká, nejsou s ní dostatečné zkušenosti, je obtížně modelovatelná. Postupy jsou v detailech pro každou aplikaci jedinečné a nejsou dostatečně formalizovány. Diskuse by se mohla týkat např. následujících problémů CBR týkajících se slabých míst přístupu CRB a možností jejich redukce a dalšího pozitivního vývoje CBR přístupů.

Jaké jsou možnosti větší formalizace přístupů?

Jaké jsou možnosti redukce např. následujících problémů:

Metodiky nejsou automatizovány.

CRB představuje mnoho aplikací, každá užívá trochu jiných metod pro uzpůsobení v rámci CBR.

Co je v rámci CBR společné, co speciální?

Existuje hierarchizace přístupů?

Existují obecné efektivní validační a verifikační metody?

## **Závěr:**

Práce splňuje v nároky na disertační práce kladené. Je nutno též konstatovat, že práce demonstruje autora již jako zralou osobnost s jistými výzkumnými a publikačními aktivitami v širokém, náročném nepřehledném, ale zajímavém oboru tematiky práce. Doporučuji, aby Ing. Filipu Mezerovi byl po úspěšné obhajobě udělen titul Ph.D.

Hradec Králové 15.5.2018



Antonín Slabý