Univerzita Pardubice
Fakulta ekonomicko-správní

Modelování pravidlových systémů v oblasti veřejné správy

Jaromír Pírko

Diplomová práce
2008

Univerzita Pardubice
Fakulta ekonomicko-správní
Ústav systémového inženýrství a informatiky
Akademický rok: **2007/2008**

ZADÁNÍ DIPLOMOVÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení:	**Bc. Jaromír PÍRKO**
Studijní program:	**N6209 Systémové inženýrství a informatika**
Studijní obor:	**Regionální a informační management - Informační management**
Název tématu:	**Modelování pravidlových systémů v oblasti veřejné správy**

Zásady pro vypracování:

Předpokládá se, že diplomová práce bude obsahovat:
- vymezení pojmů pravidlový systém, klasifikační systém, podmíněné pravidlo atd.;
- posouzení využití DBE (The Digital Business Ecosystem) v oblasti veřejné správy;
- návrh modelů pravidlového systému pro vybraný příklad z oblasti veřejné správy.

Rozsah grafických prací:

Rozsah pracovní zprávy:

Forma zpracování diplomové práce: **tištěná/elektronická**

Seznam odborné literatury:

FOTR, Jiří. Manažerské rozhodování: postupy, metody a nástroje. 1. vyd. Praha: Ekopress, 2006. ISBN 80-86929-15-9

MAŘÍK, Vladimír a kol.Umělá inteligence. 1. vyd. Praha: Academia, 1993- . ISBN 80-200-0502-1

Welcome to DBE — Digital Business Eco-System [online]. Poslední revize 12.9.2006 [cit. 2005-10-04]. Dostupné z: <http://www.digital-ecosystem.org>

WOKOUN, René. Úvod do regionálních věd a veřejné správy. 3. vyd. Praha: IFEC, 2001. ISBN 80-86412-08-3

WOKOUN, René - Mates, Pavel. Management regionální politiky a reforma veřejné správy. 1. vyd. Praha: Linde, 2006. ISBN 80-7201-547-8

WOKOUN, René. Úvod do regionálních věd a veřejné správy. 3. vyd. Praha: IFEC, 2001. ISBN 80-86412-08-3

Vedoucí diplomové práce: **doc. Ing. Jiří Křupka, Ph.D.**
Ústav systémového inženýrství a informatiky

Datum zadání diplomové práce: **12. října 2007**

Termín odevzdání diplomové práce: **26. května 2008**

prof. Ing. Jan Čápek, CSc.
děkan

L.S.

doc. Ing. Pavel Petr, Ph.D.
vedoucí ústavu

V Pardubicích dne 12. října 2007

Děkuji doc. Ing. Jiřímu Křupkovi, Ph.D., vedoucímu diplomové práce,
za vstřícnost, trpělivost a odborné vedení.

SOUHRN

Práce pojednává o využití klasifikačních pravidlových systémů jako systémů pro podporu rozhodování regionálního managementu. Obsahuje návrh dvou metod klasifikace spokojenosti občanů (rozhodovací stromy a fuzzy inferenční systémy), návrh modelů spokojenosti občanů se životním prostředím v Chrudimi a hodnocení jejich výsledků.

KLÍČOVÁ SLOVA

klasifikační systémy; fuzzy inferenční systémy; rozhodovací stromy; regionální management; Společné evropské indikátory; spokojenost občanů

TITLE

Rules Systems Modelling in the Public Administration Area

ABSTRACT

This work deals with using of the rules classification systems as the regional management decision support systems. It includes the design of two citizen satisfaction classification methods (decision trees and fuzzy inference systems), the design of models of the citizen satisfaction with Chrudim environment and evaluation of its results.

KEYWORDS

classification systems; fuzzy inference systems; decision trees; regional management; The European Common Indicators; citizen satisfaction

Obsah

Seznam obrázků 7
Seznam tabulek 7
Seznam použitých zkratek 8
Úvod 8
1 Regionální management ve veřejné správě 10
1.1 Management - systém řízení 10
1.2 Management - řídící systém 12
1.3 Specifika regionálního managementu 14
1.4 Informace o spokojenosti občanů - zpětná vazba regionálního managementu 18
1.4.1 Hodnocení regionálního managementu 20
1.4.2 Přímé měření spokojenosti občanů 22
1.4.3 Současné postupy zpracování šetření spokojenosti 23
2 Návrh úloh zpracování dat o spokojenosti občanů 25
2.1 Iniciativa Společné evropské indikátory 25
2.2 Dotazníkový průzkum Spokojenost občanů s místním společenstvím v Chrudimi 27
3 Metody využitelné pro řešení úloh o spokojenosti občanů 29
3.1 Pravidlové systémy 29
3.2 Klasifikační systémy založené na pravidlech 30
3.3 Formalizace klasifikační úlohy 33
3.4 Rozhodovací stromy 35
3.5 Fuzzy inferenční systémy 36
3.5.1 Fuzzy množiny a fuzzy usuzování 37
3.5.2 Struktura fuzzy inferenčních systémů 38
3.6 Návrh metod řešení klasifikační úlohy o spokojenosti občanů 40
4 Využití Digitálního ekosystému pro podnikání v oblasti veřejné správy 41
5 Návrh modelů klasifikačního systému 44
5.1 Formalizace klasifikační úlohy a návrh atributů 44
5.2 Popis dat 47
5.3 Příprava dat 48
5.4 Modelování spokojenosti - rozhodovací stromy 49
5.5 Modelování spokojenosti - fuzzy inferenční systémy 53
5.6 Vyhodnocení výsledků 61
5.6.1 Vyhodnocení modelů rozhodovacích stromů 63
5.6.2 Vyhodnocení fuzzy modelů 66
6 Závěr 69
7 Literatura 72
Příloha 1: Dotazník – Indikátor 1 77
Příloha 2: Metodika dotazníkového průzkumu 79

Seznam obrázků

Obr. 1 Řízení bez zpětné vazby, podle ŠVARC (2002) 11
Obr. 2 Řízení se zpětnou vazbou, podle ŠVARC (2002) 11
Obr. 3 Řízení organizace 12
Obr. 4 Model 7-S, upraveno podle RECKLIES (2008) 13
Obr. 5 Řízení výrobního podniku, upraveno podle KOŠČOVÁ - EXNAR - KŘUPKA (1997) 15
Obr. 6 Řízení regionu, upraveno podle KOŠČOVÁ - EXNAR - KŘUPKA (1997) a PŮČEK (2005) 15
Obr. 7 Systém řízení podniku a jeho okolí, upraveno podle TURBAN – ARONSON – LIANG (2005) 17
Obr. 8 Systém řízení regionu a jeho okolí, upraveno podle TURBAN – ARONSON – LIANG (2005) a PŮČEK (2005) 17
Obr. 9 Schéma rozhodovacího procesu se zpětnou vazbou, upraveno podle KOORY – MEDLEY (1987) 19
Obr. 10 Projekt a cyklus šetření spokojenosti, PŮČEK (2005) 22
Obr. 11 Model hodnocení spokojenosti občanů jako subsystém systému řízení 26
Obr. 12 Klasifikátor, KOTEK (1980) 32
Obr. 13 Operace mezi fuzzy množinami, THE MATHWORKS (1998) 38
Obr. 14 Model hodnocení spokojenosti občanů 40
Obr. 15 Čtyři hlavní rysy DBE, ERISA (2007) 41
Obr. 16 DBE jako model pro místní rozvoj, ERISA (2007) 42
Obr. 17 Histogram cílového atributu 50
Obr. 18 Stream pro tvorbu modelu DT1 51
Obr. 19 Náhled části rozhodovacího stromu modelu DT1 52
Obr. 20 Rozdělení dat a úprava vybraných atributů v Clementine 55
Obr. 21 Model DT1 - porovnání křivek navýšení 64
Obr. 22 Model DT4 - porovnání křivek navýšení 65

Seznam tabulek

Tab. 1 Základní tvar fuzzy pravidlového systému, upraveno podle ROSS (2004) 38
Tab. 2 Přehled vybraných atributů 46
Tab. 3 Pearsonův koeficient korelace 47
Tab. 4 Datový slovník 48
Tab. 5 Přehled modelů a jejich vstupních atributů 56
Tab. 6 Matice záměn, podle BERKA (2003). 62
Tab. 7 Model DT1 - matice záměn 63
Tab. 8 Model DT2 - matice záměn 64
Tab. 9 Model DT3 - matice záměn 64
Tab. 10 Model DT4 - matice záměn 65
Tab. 11 Model DT5 - matice záměn 65
Tab. 12 Model DT6 - matice záměn 65
Tab. 13 Model DT7 - matice záměn 66
Tab. 14 Model FIS1 - matice záměn 66
Tab. 15 Model FIS2 - matice záměn 66
Tab. 16 Model FIS3 - matice záměn 67
Tab. 17 Model FIS4 - matice záměn 67
Tab. 18 Model FIS5 - matice záměn 67
Tab. 19 Přehled výsledků klasifikace 68

Seznam použitých zkratek

DBE	The Digital Business Ecosystem (Digitální ekosystém pro podnikání)
ECI	The European Common Indicators (Společné evropské indikátory)
ERISA	The European Regional Information Society Association
FIS	Fuzzy Inference System (fuzzy inferenční systém)
mf	membership function (funkce příslušnosti)
MSP	malé a středí podniky
NKÚ	Nejvyšší kontrolní úřad
NPVP	Národní plán vyrovnávání příležitostí
PDCA	Plan Do Check Act (Plánuj, proveď, ověř, proveď opatření)
SUR	Strategie udržitelného rozvoje České republiky
TIMUR	Týmová iniciativa pro místní udržitelný rozvoj

Úvod

Spokojenost člověka je jednou ze základních podmínek určujících kvalitu jeho života. Je to však veličina značně subjektivní a měnící se v čase a sám pojem spokojenost je rozsáhlý a neurčitý. Mezi základní cíle regionálního managementu patří rozvoj regionu a změna kvality života jeho občanů. Působení na zlepšování kvality života občanů a tím i na zvyšování míry jejich spokojenosti vyžaduje, aby byly k dispozici nástroje, které umožní posoudit úspěch tohoto působení. Informace o spokojenosti občanů jsou pro regionální management významným podkladem pro rozhodování a sebehodnocení, proto je třeba spokojenost občanů nějak měřit a hodnotit.

V současné době využívá regionální management k měření spokojenosti občanů převážně dotazníková šetření týkající se spokojenosti občanů s různými aspekty života. Výsledky jsou obvykle zpracovávány pomocí běžných statistických metod. Každý dotazovaný aspekt spokojenosti je při nich vyhodnocován samostatně. Toto zpracování je buď prezentováno jako konečný výstup, někdy doplněný komentářem, nebo je podkladem pro následné sociologické analýzy. Tyto postupy jsou z hlediska zpracování poměrně jednoduché. Mají své nesporné výhody, ale nevyužívají dostatečně možnosti, které získaná data nabízejí. Při šetřeních jsou shromažďovány velké objemy dat s významným informačním potenciálem, nad kterými je možné formulovat různé typy úloh. Nabízí se zde tedy široké pole působnosti pro jejich využití pomocí umělé a výpočetní inteligence.

Jednou z možností, jak může regionální management využít data získaná dotazníkovým šetřením spokojenosti občanů je hledání souvislostí mezi demografickými údaji o občanech, dílčími aspekty spokojenosti s životem v obci, aspekty spokojenosti s různými rysy samosprávy a celkovou spokojeností občanů. Pokud by se podařilo definovat skupiny občanů, kteří jsou nespokojeni, můžou získané znalosti sloužit jako podklady pro nalezení příčin tohoto stavu a pro stanovení priorit při dalším rozhodování.

Z hlediska dobývání znalostí lze úlohu formulovat např. jako nalezení takových socio-demografických údajů o občanech a takových aspektů spokojenosti občanů, které umožní určit míru jejich spokojenosti s některým z dílčích aspektů, popř. určit míru jejich celkové spokojenosti. Úloha proto může být pojata jako úloha klasifikační - cílem je zařazení občanů do tříd podle míry jejich spokojenosti. Velký význam by pro regionální management měla možnost klasifikace občanů na základě socio-demografických dat, která jsou již obsažena v databázích městských úřadů a registrech občanů a není je tedy nutné získávat nákladným,

časově náročným a méně přesným dotazníkovým šetřením. Tato data jsou k dispozici okamžitě, navíc jsou průběžně aktualizována.

Hlavním cílem diplomové práce je návrh metod vhodných pro klasifikaci občanů do tříd podle míry jejich spokojenosti. Dílčí cíle představují:

- rešerše odborné literatury související s problematikou veřejné správy a literatury z oblasti klasifikačních systémů založených na pravidlech;
- popis specifických rysů hodnocení spokojenosti občanů a definování klasifikační úlohy o spokojenosti občanů;
- předzpracování reálných výsledků dotazníkového šetření a realizace vybraných metod s podporou vhodných softwarových nástrojů;
- návrh a hodnocení modelů klasifikace spokojenosti a porovnání zvolených metod.

V diplomové práci je popsán management jako vědecká disciplína, řídící systém i systém řízení. Dále jsou uvedena specifika regionálního managementu, popsáno je jeho postavení z pohledu teorie managementu a vymezena jeho role ve struktuře veřejné správy. Další část je věnována významu informací o spokojenosti občanů jako zpětné vazby regionálního managementu, včetně popisu současných postupů hodnocení spokojenosti.

Druhá kapitola obsahuje návrhy úloh zpracování dat o spokojenosti občanů a uvádí možné zdroje dat, zejména výsledky dotazníkových průzkumů. Třetí kapitola popisuje metody řešení úloh o spokojenosti občanů se zaměřením na formalizaci klasifikační úlohy a její řešení pomocí pravidlových systémů. Následuje posouzení využití Digitálního ekosystému pro podnikání v oblasti veřejné správy a možnosti zahrnutí modelů hodnocení spokojenosti občanů do jeho struktur. Pátá kapitola obsahuje návrhy a realizaci modelů pravidlového systému klasifikace občanů podle jejich spokojenosti se životním prostředím v Chrudimi, včetně popisu a přípravy vstupních dat získaných dotazníkovým průzkumem Spokojenost občanů s místním společenstvím v Chrudimi. Modely s nejlepšími výsledky jsou popsány, vyhodnoceny a porovnány. Na závěr jsou shrnuty problémy, kterým by bylo účelné věnovat pozornost.

Pro klasifikaci občanů byly použity dvě metody založené na rozhodovacích pravidlech - algoritmy pro tvorbu rozhodovacích stromů a fuzzy inferenční systémy. Pro modelování a vyhodnocení modelů je využito prostředí SPSS Clementine a MATLAB.

1 Regionální management ve veřejné správě

Pojem management má v angličtině i v češtině více významů. Řízení nebo vedení může znamenat soubor řídících znalostí a dovedností a vědní disciplínu, která se jimi zabývá (teorie řízení[1]). Znamená však hlavně posloupnost cílevědomých činností (systém řízení) a odborníky v čele organizačních jednotek (řídící systém). Tato kapitola je věnována managementu v druhém a třetím významu se zaměřením na management územní samosprávy - regionální management.

1.1 Management - systém řízení

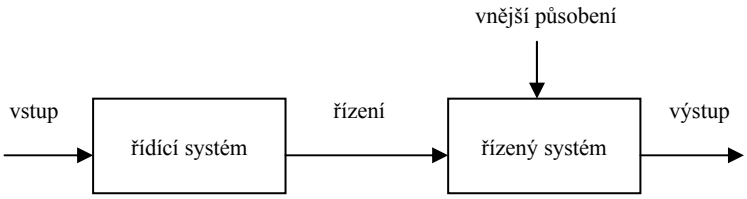
Management ve významu posloupnost činností definují TURBAN - MEREDITH (1991) stručně jako „proces dosahování určitých cílů používáním zdrojů (lidí, peněz, energie, materiálu, prostoru, času); zdroje jsou považovány za vstupy, dosažení cílů za výstup procesu“. Tentýž pojem uvádí KRATOCHVÍL (2000) v pojetí praktických manažerů jako „tvůrčí a odpovědnou činnost, směřující k dosažení trvale se vyvíjející množiny cílů, prováděnou v souladu se změnami řízeného objektu, řídícího subjektu i jejich okolí“. Obě definice shodně využívají systémový přístup - považují proces řízení implicitně za systém. Ten je vymezen, ŠTACH (1982), jako množina prvků a vazeb, která má jako celek určité vlastnosti. Protože existují vazby mezi systémem a jeho okolím (vstupy a výstupy), je řízení systémem otevřeným.

Důležitým hlediskem pro rozlišení řízení podle jeho účinnosti je možnost kontroly, zda výsledek řízení je anebo není zpětně kontrolován, zda je či není zpětná vazba při řízení. ŠVARC (2002). Podle toho rozlišujeme ovládání, regulaci a vyšší formy řízení:

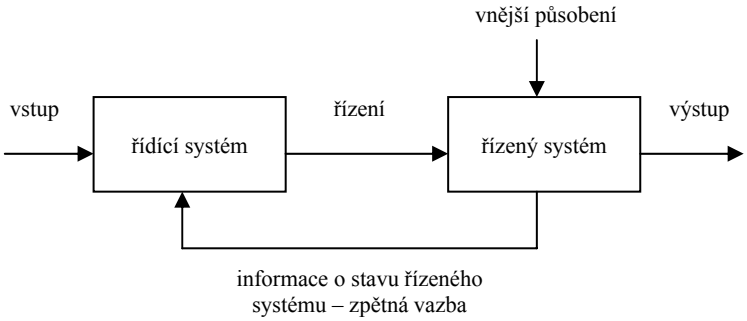
- ovládání je řízení bez zpětné kontroly - bez zpětné vazby (obr. 1);
- regulace je řízení se zpětnou vazbou (obr. 2);
- vyšší formy řízení - optimální řízení, adaptivní řízení, učení a umělá inteligence.

V dalším textu bude pojmu řízení užíváno ve významu řízení se zpětnou vazbou.

[1] V případě managementu jako vědní disciplíny je vhodné rozlišovat teorii automatického řízení, čímž zdůrazňujeme, že se jedná o řízení technických zařízení (Automatic Control), protože řízení společenské (Management) se spíš označuje bez přívlastku pouze jako teorie řízení. ŠVARC (2002)



Obr. 1 Řízení bez zpětné vazby, podle ŠVARC (2002)



Obr. 2 Řízení se zpětnou vazbou, podle ŠVARC (2002)

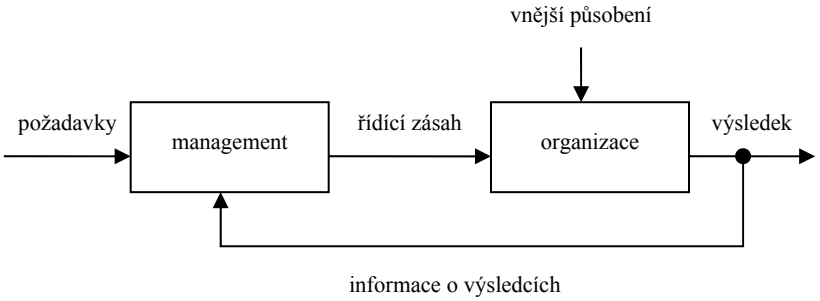
Řízení se zpětnou vazbou je udržováním určité veličiny na konstantní hodnotě nebo na hodnotě měnící se podle určitého pravidla. Pokud je zjištěna odchylka mezi požadovanou hodnotou výstupu (cílem) a výstupem systému, řídící systém provede řídící zásah, kterým je odchylka odstraněna. Odchylka může být způsobena špatným nastavením řídícího systému nebo vnějším působením.

Zpětná vazba tvořící uzavřenou smyčku zajišťuje tok informací o výstupu systému zpět do řídícího systému. Na jejím základě mění řídící systém vstupy, procesy[2] nebo obojí, tak, aby se přiblížil očekávanému cíli. TURBAN – ARONSON – LIANG (2005). Zpětná vazba je tedy podstatou řízení a v této práci má zásadní význam.

[2] Procesy jsou všechny prvky nezbytné k přeměně vstupů na výstupy.

1.2 Management - řídící systém

Management jako řídící systém je management ve významu týmu odborníků na řízení - manažerů**,** kteří stojí v čele organizační jednotky a kteří jsou zodpovědní za dosahování jejích cílů. K dosažení těchto cílů je nezbytné, pomocí určitých metod a technik, zvládnout specifické činnosti - manažerské funkce. Management je výše zmíněným řídícím systémem a je tedy subsystémem systému řízení (obr. 3).



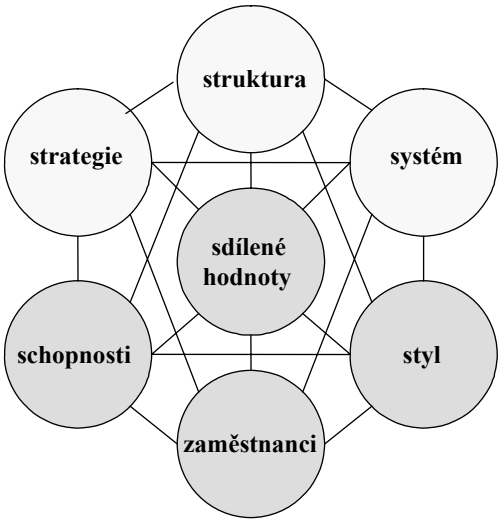
Obr. 3 Řízení organizace

Úspěch managementu, TURBAN - MEREDITH (1991), závisí na vykonávání manažerských funkcí, kterými jsou plánování, organizování, vedení a kontrola.

Detailnější rozpracování manažerských funkcí na sekvenční a paralelní uvádí jeden z předních představitelů světového managementu Peter Ferdinand Drucker. Dosažení cílů organizace může být nejlépe zajištěno harmonií všech těchto funkcí. SIEGL (2002)[3]. Manažerské funkce sekvenční jsou plánování, organizování, výběr a rozmístění spolupracovníků, vedení lidí a kontrola. Manažerské funkce paralelní jsou analyzování řešených problémů, rozhodování a implementace rozhodnutí.

[3] Autor cituje z knihy DRUCKER, Peter F. The Practice of Management. New York: Harper and Raw, 1952.

Příkladem manažerské metody může být model 7-S (obr. 4), který je uplatňován od osmdesátých let. Specifikuje oblasti, které je třeba z pohledu managementu identifikovat, a jejich společným nasměrováním účinněji dosahovat cílů. Změní-li se jeden z prvků modelu, ovlivní to i prvky ostatní.



Obr. 4 Model 7-S, upraveno podle RECKLIES (2008)

Model[4] rozlišuje tvrdá S (světlá pole) a měkká S (tmavší pole). Měkká S jsou sice obtížněji identifikovatelná, ale nejúspěšnější společnosti intenzivně pracují právě na nich.

Tvrdá S:

- strategie (Strategy) - dlouhodobé záměry;
- struktura (Structure) - základ pro specializaci a uspořádání ovlivněný v první řadě strategií, velikostí a rozmanitostí;
- systém (System) - formální a neformální procedury, které podporují strategii a strukturu.

Měkká S:

- zaměstnanci (Staffs) - lidé a řízení lidských zdrojů;
- sdílené hodnoty (Shared Values) - základní myšlenky, na kterých organizace stojí;
- styl (Style) - organizační kultura, chování manažerů;
- schopnosti (Skills) - charakteristické schopnosti, co dělá organizace nejlépe.

[4] Model vyvinuli Tom Peters a Robert Waterman z McKinsey & Co, proto se užívá i označení McKinsey 7-S. Publikován byl poprvé v článku "Structure Is Not Organization" v roce 1980.

1.3 Specifika regionálního managementu

Jedním z možných pohledů na uspořádání veřejné správy v České republice je její hierarchické členění na základní části - státní správa, územní samospráva a subjekty pověřené výkonem veřejné správy (veřejné fondy a nadace, zmocněné právnické a fyzické osoby).

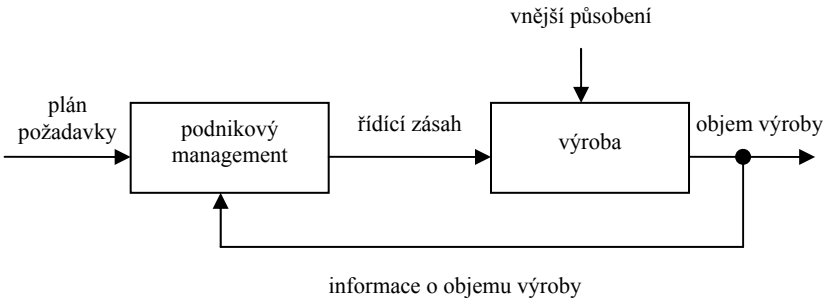
Z hlediska územního členíme veřejnou správu na ústřední (správa území celého státu), oblastní (správa jednotlivých krajů, dříve i okresů), místní (správa jednotlivých obcí).

Správní členění vychází z Ústavy České republiky. Ústavou je zaručena také samospráva územních samosprávných celků. To podtrhuje jejich význam jako autonomních organizačních jednotek s potřebou vlastního řízení. Důraz na samosprávné postavení klade DRUCKER (2006): „Alternativou vůči samosprávným institucím, které fungují a podávají výkon, není svoboda. Jejich alternativou je totalitní hrůzovláda."

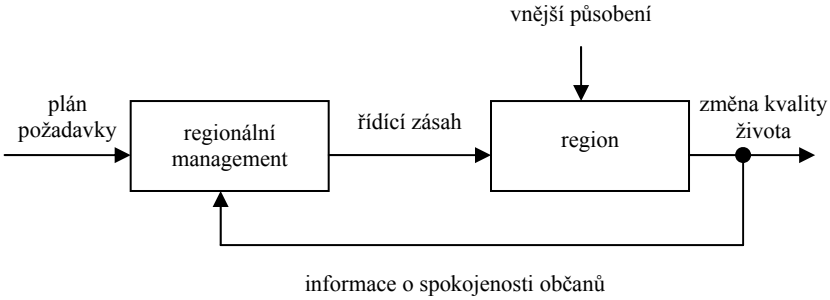
Tato práce je zaměřena právě na management regionální, tzn. management územní samosprávy na místní úrovni (zde management obce). Pojmy řízení (Management) a správa (Administration) nejsou synonyma, stejně tak jimi není ani jejich užívání ve spojení management veřejné správy a veřejná správa. Podle WRIGHT (2003) byla veřejná správa uznávaným pojmem pro studium řízení veřejného sektoru během téměř celého dvacátého století. Během osmdesátých let došlo v Anglii a ve Spojených státech k vývoji, který podnítil myšlenku, že veřejná správa zcela neobsahuje to, co je požadováno od manažerů ve veřejné správě na rozdíl od veřejných úředníků. V případě pracovníků ve veřejné správě jde zejména o vykonávání instrukcí, management se vyznačuje, jak již bylo uvedeno, zodpovědností za formulaci a dosahování cílů.

Regionální management definuje ADAMČÍK (2000) jako „soubor přístupů, názorů, zkušeností, doporučení a metod, které vedoucí pracovníci regionů, obcí a měst (regionální manažeři) užívají nebo by měli využívat ke zvládnutí specifických činností, jež jsou potřebné k dosažení dvou vzájemně závislých cílů regionální politiky: ekonomického růstu regionů, obcí a měst a ke zlepšování prostorového rozdělení". Tato definice je v podstatě analogií k definicím podnikového managementu a je zaměřena na ekonomickou stránku problému. Sám autor její část věnovanou cílům na jiném místě zobecňuje - „cílem regionálního managementu je prospěch a rozvoj regionu, uspokojování zájmů a potřeb obyvatelstva a jeho skupin a veřejný zájem". Stejně jej definuje i zákon č. 128/2000 Sb., o obcích (obecní zřízení) a zákon č. 129/2000 Sb., o krajích (krajské zřízení).

Systém řízení organizace v soukromém sektoru je v základních prvcích a při malé rozlišovací úrovni obdobný jako systém řízení ve veřejné správě. Porovnání obou systémů na příkladech reálných systémů řízení výrobního podniku a regionu je na obr. 5 a na obr. 6.



Obr. 5 Řízení výrobního podniku, upraveno podle KOŠČOVÁ - EXNAR - KŘUPKA (1997)



Obr. 6 Řízení regionu, upraveno podle KOŠČOVÁ - EXNAR - KŘUPKA (1997) a PŮČEK (2005)

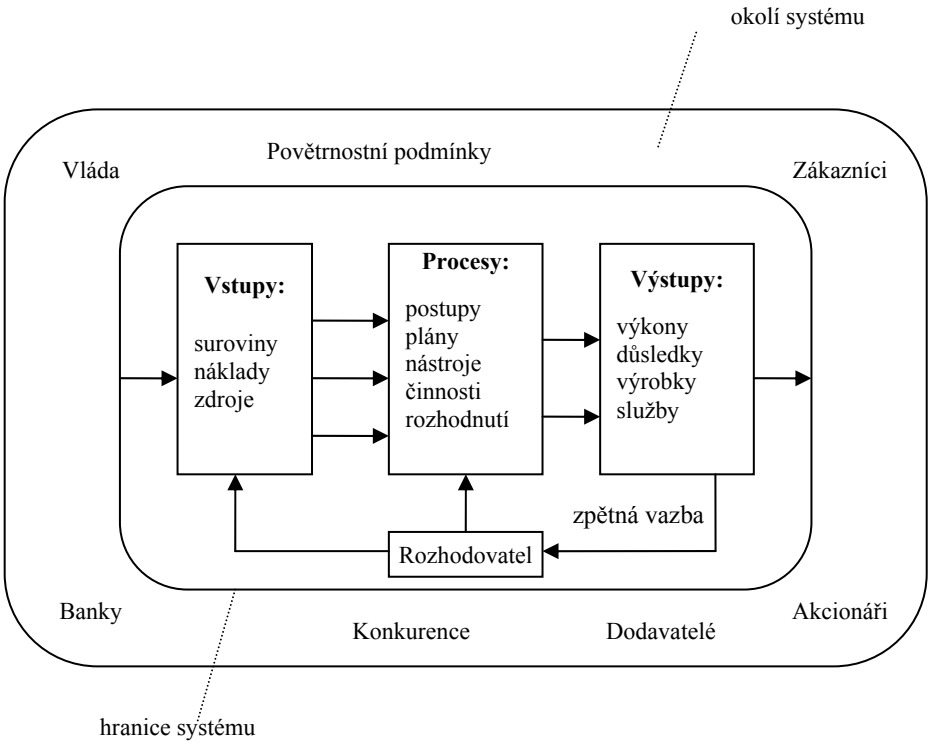
K první systematické aplikaci teorie a principů managementu nedošlo v podnikové organizaci. První cílevědomou a systematickou aplikací principů managementu ve Spojených státech - prováděnou s plným vědomím toho, že jde o aplikaci managementu - byla podle DRUCKER (1992) reorganizace americké armády. O několik let později, v roce 1908, se objevil první městský manažer (Staunton, VA), což bylo výsledkem aplikace zcela nových principů managementu jako oddělení „politiky“ (svěřené volené a politicky odpovědné městské radě) od managementu (svěřeného nepolitickému profesionálovi, odpovědnému po linii řízení). Nikdo již nemohl považovat ani nové podnikové organizace za přímé pokračování staré a tradiční „obchodní firmy“ - „kanceláře“, již tvořili dva postarší bratři a jeden úředník a která figuruje v knihách Charlese Dickense.

Management veřejné správy má tedy své specifické rysy a vlastnosti, které nejsou jen prostou aplikací ekonomického či podnikového managementu. Důvody všeobecné představy o dominanci podnikového managementu vysvětluje DRUCKER (2006): „Podnik není jakousi vzácnou výjimkou - je prostě první z druhů, kterým jsme až dosud věnovali nejdůkladnější pozornost.“ Existují mnohé významné rozdíly, které uvádějí ADAMČÍK (2000), GROSPIČ (2006), LUKÁŠ (2000), WRIGHT (2003):

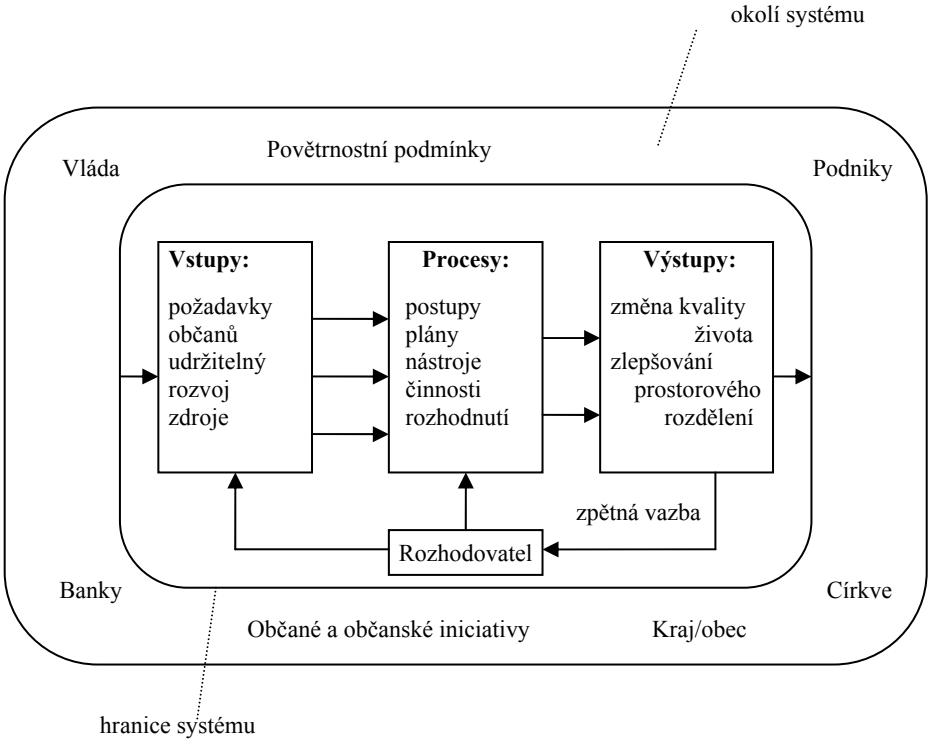
- cílem podnikového managementu je maximalizace zisku, cílem regionálního managementu je prospěch a rozvoj regionu;
- základem měření výkonnosti v soukromém sektoru je zisk, což ve veřejném sektoru nelze použít;
- analogicky ke správnímu právu lze použít zásadu, že manažer v soukromém sektoru smí vše, co není zakázáno, kdežto manažer ve veřejném sektoru smí jen to, co mu ukládá zákon;
- veřejný zájem je často odlišný od soukromých zájmů;
- pracovníci veřejné správy jsou odpovědni demokratickým hodnotám;
- ústava požaduje rovné zacházení s jednotlivci a vylučuje druh selektivity, která je často podstatná k udržení zisku;
- zatímco v ekonomickém systému jsou řízeným systémem přímo reálné produkční entity, v případě veřejné správy jimi jsou pouze vztahy mezi relevantními entitami reálného světa;
- mnoho rozhodnutí ve veřejném sektoru musí být veřejných, v soukromém sektoru nemusí a mnohdy by to mohlo podnik i ohrozit.

V posledních desetiletích byly vytvořeny i specifické nástroje regionálního managementu. Jako hlavní skupiny nástrojů uvádí ADAMČÍK (2000) informační opatření a poradenství, finanční nástroje, opatření v infrastruktuře a regulační opatření. Vzhledem k zadání této práce je zajímavá hlavně skupina první. Více místa je jí věnováno ve 4. kapitole Digitální ekosystém pro podnikání.

Porovnání systémů řízení podniku a regionu, stejných jako na výše uvedených obr. 5 a obr. 6, ale při větší úrovni rozlišení, je na obr. 7 a 8. Vidíme, že řídící systém je obdobný, ale liší se vstupy, výstupy i relevantní okolí .



Obr. 7 Systém řízení podniku a jeho okolí, upraveno podle TURBAN – ARONSON – LIANG (2005)



Obr. 8 Systém řízení regionu a jeho okolí, upraveno podle TURBAN – ARONSON – LIANG (2005) a PŮČEK (2005)

Mezi podnikovým managementem a managementem veřejné správy sice existují významné rozdíly už z jejich podstaty, ale z pohledu uplatňování základních manažerských funkcí zásadní rozdíly nejsou, naopak najdeme více podobností. Proto lze získané zkušenosti vzájemně uplatňovat. Najdeme mnoho příkladů, kdy byla některá metoda uplatněna v soukromém sektoru a poté převzata manažery veřejné správy. Jsou to např. metody:

- ZBB - rozpočtování s nulovou bází (Zero-based Budgeting),
- TQM - management kvality (Total Quality Management),
- MBO - cílovaný management (Management by Objectives) aj.

Tak jako dochází ke stírání rozdílů mezi managementem soukromého a veřejného sektoru, dochází současně i ke sbližování sektorů samotných. Je patrný odklon od čistě soukromých a čistě veřejných organizací. Existují organizace částečně vlastněné státem (aerolinie, dopravní podniky), regulované soukromé organizace (rozvody vody, telekomunikace), ale i veřejná infrastruktura spravovaná soukromým subjektem na principu „postav - provozuj - převeď“ (dálnice, věznice).

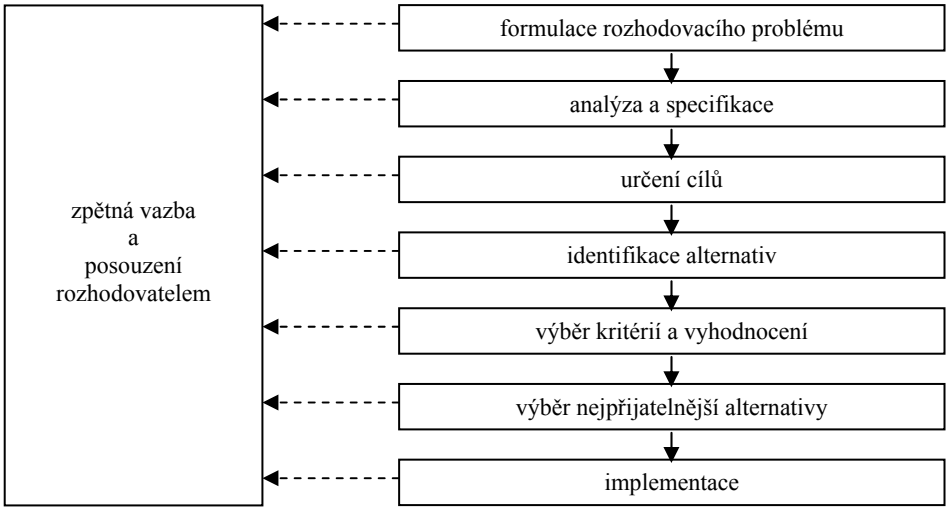
Dalším příkladem sbližování je spolupráce označovaná jako partnerství veřejného a soukromého sektoru (PPP - Public Private Partnerships). PPP je obecný pojem pro spolupráci veřejného a soukromého sektoru vzniklou za účelem využití zdrojů a schopností soukromého sektoru při zajištění veřejné infrastruktury nebo veřejných služeb. Jednotlivé varianty PPP, jsou-li odborně a úspěšně aplikovány, zvyšují kvalitu i efektivnost veřejných služeb včetně výkonu státní správy a urychlují realizaci významných infrastrukturních projektů s pozitivním dopadem na rozvoj ekonomiky. PPP CENTRUM (2008). V devadesátých letech 20. století se objevuje koncept New Public Management - model řízení organizace veřejné správy. Je to řízení orientované na službu, kdy řízení probíhá prostřednictvím stanovení cílů a delegování výkonných pravomocí na regionální úroveň. Řečeno slovy P. F. Druckera: „Management se stal všudypřítomným a univerzálním nástrojem moderní společnosti.“ DRUCKER(1992).

1.4 Informace o spokojenosti občanů - zpětná vazba regionálního managementu

V předchozích podkapitolách je popsán význam informací o výstupu systému, které putují zpět k řídícímu systému - managementu organizace a které umožňují jeho porovnání

s očekávaným cílem. Tyto informace mohou mít různou podobu, využívají různé informační kanály a podle účelu jejich dalšího využití se může lišit i způsob jejich interpretace.

Pro regionální management jsou jedním z nejdůležitějších druhů informací zpětné vazby informace o spokojenosti občanů s různými rysy územní samosprávy. Zásadní význam mají zejména pro vykonávání sekvenční manažerské funkce plánování a paralelních manažerských funkcí analyzování řešených problémů a rozhodování. Příkladem významu zpětné vazby pro rozhodování je krokový model všeobecného rozhodovacího procesu se zpětnou vazbou (obr. 9). Každý krok je zde konfrontován čárkovaně znázorněnými zpětnými vazbami vůči celému řídícímu procesu. Svůj význam má tento druh informací i pro výkon ostatních manažerských funkcí.



Obr. 9 Schéma rozhodovacího procesu se zpětnou vazbou, upraveno podle KOORY – MEDLEY (1987)

Povinnost regionálního managementu starat se o spokojenost občanů vyplývá ze zákona č. 128/2000 Sb., o obcích (obecní zřízení): „Obec pečuje o všestranný rozvoj svého území a o potřeby svých občanů; při plnění svých úkolů chrání též veřejný zájem.“ Stejná povinnost vyplývá i ze zákona č. 129/2000 Sb., o krajích (krajské zřízení): „Kraj pečuje o všestranný rozvoj svého území a o potřeby svých občanů.“

ČSN EN ISO 9000:2001, Systémy managementu jakosti – Základy, zásady a slovník, ČSN (2002), definuje termín spokojenost zákazníka jako „vnímání zákazníka týkající se stupně splnění jeho požadavků“. V oblasti územní samosprávy je zákazníkem regionálního managementu občan, poskytovatelem služeb je regionální management.

Informace zpětné vazby o spokojenosti občanů s různými rysy samosprávy jsou pro regionální management především významnými podklady pro:

- rozhodování o přesunu uspokojování potřeb do místa jejich vzniku;
- rozhodování o cílových skupinách;
- demonstraci transparentnosti využití veřejných zdrojů (např. směrování investic nebo přidělování dotací);
- dokumentaci výkonnosti managementu;
- regionální marketing (produktem je celý region jako sociální systém);
- analýzu aglomeračních výhod a nevýhod;
- formulaci regionální koncepce trvale udržitelného rozvoje.

Tyto informace napomáhají také při hledání:

- co nejúčinnější organizace chodu úřadu a organizací řízených orgány samosprávy;
- způsobů zlepšování image úřadu i samotného regionu;
- alternativních řešení pro definování vize, strategie, plánů;
- ověření správnosti vize, strategie, plánů.

ADAMČÍK (2000), LUKÁŠ (2000)

Pro získání informací o spokojenosti občanů lze v současnosti využít dvou velkých skupin metod a nástrojů, z nichž každá hodnotí jiný prvek systému řízení. Nejsou v rozporu, naopak jsou běžně využívány společně. Jeden možný přístup je hodnocení řídícího systému - managementu. Z jeho výsledků lze dělat závěry o výstupu. Druhou možností je přímé měření dopadů výstupu systému - spokojenosti občanů.

1.4.1 Hodnocení regionálního managementu

Ministerstvo vnitra podporuje zavádění nástrojů řízení kvality ve veřejné správě mj. i prostřednictvím metodického vedení a koordinace. Rada České republiky pro jakost zpracovala návrh nové Strategie Národní politiky kvality v ČR na období let 2008 až 2013. Návrh Strategie vychází z analýzy výsledků dosavadního plnění Národní politiky podpory jakosti a na základě vyhodnocení současné situace definuje na příští období poslání, vizi, rámec a dlouhodobý strategický cíl. Tím je spoluvytvářet v ČR prostředí, ve kterém je úsilí o vysokou kvalitu trvalou součástí všech oblastí života společnosti i občanů a růst kvality života v České republice je veden cestou udržitelného rozvoje. VORLÍČEK (2008). Strategie má být

přínosem pro podnikatelskou sféru, veřejnou správu i pro celou občanskou společnost. V oblasti veřejného sektoru má Strategie podporovat rozvoj kvality veřejné správy.

K této problematice se v lednu 2008 v Karlových Varech uskutečnila 4. Národní konference kvality ve veřejné správě, kterou uspořádalo Ministerstvo vnitra a Karlovarský kraj. V říjnu 2008 se v Paříži uskuteční 5. konference kvality. Hlavním tématem konference se stává „Občan v srdci kvalitní veřejné správy". Organizátoři popisují téma konference takto JALOVECKÁ (2008): „Hlavním cílem veřejné správy je zlepšení života občanů. Abychom k tomuto cíli dospěli, je nezbytné zapojit do veřejné správy řízení procesů kvality." Nejčastěji používané nástroje, metody a modely pro hodnocení kvality managementu veřejné správy jsou uvedeny v následujících odstavcích.

Normy ČSN EN ISO 9000:2001, Systémy managementu jakosti – Základy, zásady a slovník, ČSN EN ISO 9004:2001, Systémy managementu jakosti – Směrnice pro zlepšování výkonnosti a další.

Model CAF (Common Assessment Framework), Společný hodnotící rámec, byl zpracován Evropským institutem pro veřejnou správu jako nástroj pro zvyšování kvality organizací veřejné správy v Evropě.

EFQM (European Foundation for Quality Management), Evropská nadace pro management kvality - posláním je růst konkurenceschopnosti a efektivnosti evropských organizací bez ohledu na velikost a sektor působení.

Benchmarking (porovnání výkonnosti) - vzájemné srovnávání, zvyšování výkonu, kvality a učení se jeden od druhého

PDCA (Plan Do Check Act), Demingův cyklus Plánuj, proveď, ověř, proveď opatření - cyklus zlepšování práce managementu.

BSC (Balanced Scorecard), Metoda vyvážených ukazatelů.

Reengineering - přehodnocení a rekonstrukce procesů organizace veřejné správy

Charta občana - veřejný dokument, kterým se stanoví základní informace o poskytovaných službách, o úrovni služeb, kterou může zákazník očekávat, stejně tak jako způsob podávání stížností či návrhů na zlepšení.

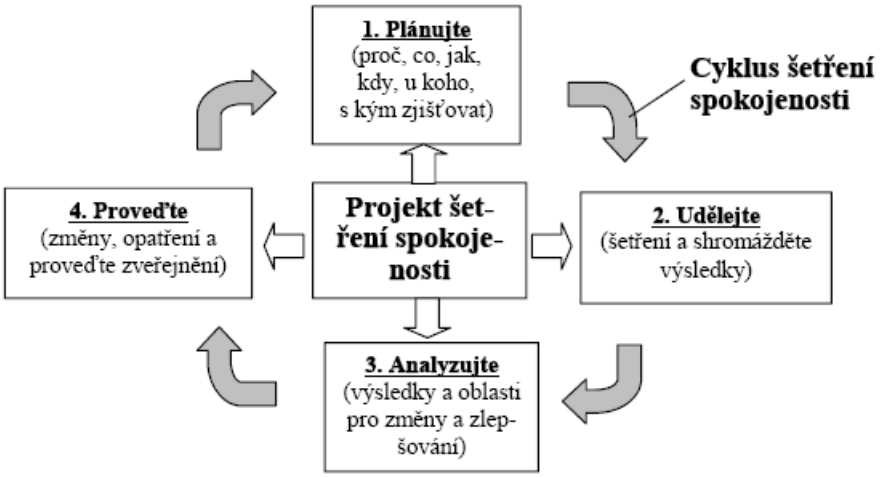
Další používané postupy nejrůznějšího charakteru a uplatnění, jako např. procesní přístup, komunitní plánování, Místní Agenda 21, dobrá praxe, externí či interní audity.

MINISTERSTVO VNITRA (2005), PŮČEK (2005)

1.4.2 Přímé měření spokojenosti občanů

Tento způsob zjišťování poskytuje managementu informace o názorech občanů přímo od občanů samotných. Přímé měření má své nesporné výhody, ale přináší i některé problémy. Ačkoliv je sám pojem „spokojenost“ definován dokonce normou ČSN EN ISO 9000:2001 (viz úvod této kapitoly), jeho nejednoznačnost je zřejmá. Velmi podobný problém, v uvedeném případě s pojmem „image“ při výzkumu „Image obce - pohled občana“, popisují BRYCHTOVÁ - DUPLINSKÝ (2005). Při rigorózním pojetí pojmu vystupuje do určité míry jeho vágnost v kognici občanů. Mnozí si jej nespojovali s městem jako místem jejich života, ale odrážel se v jejich odpovědích výrazně i vztah k vedení města a jeho představitelům.

Přes tuto pojmovou nejednoznačnost jsou informace získané měřením spokojenosti občanů velice důležitou a mnohdy jedinou zpětnou vazbou managementu veřejné správy. Výzkumy, průzkumy a šetření týkající se spokojenosti občanů s různými stránkami života jsou v praxi často užívaným nástrojem získávání informací. Provádějí se průzkumy spokojenosti občanů s osobním životem, životem v obci, fungováním samosprávy, politickou situací, bezpečností, životním prostředím, kulturou v obci atd. Jeden z možných modelů šetření - model Projektu a cyklu šetření spokojenosti, který vychází z obecného cyklu zlepšování PDCA (viz podkapitola 1.4.1), je znázorněn na obr. 10.



Obr. 10 Projekt a cyklus šetření spokojenosti, PŮČEK (2005)

Na evropské, národní i regionální úrovni provádějí šetření spokojenosti Eurobarometr (Evropská komise), Český statistický úřad, Centrum pro výzkum veřejného mínění (Sociologický ústav AV ČR), agentury pro výzkum veřejného mínění, média i samotné orgány státní správy a územní samosprávy. V praxi městských úřadů se provádějí nejčastěji obecná šetření spokojenosti občanů s životem v obci a šetření spokojenosti „zákazníků“ městského úřadu

s jeho službami. Častější je tato praxe pochopitelně ve větších městech. Povinnost provádět tato šetření není explicitně vymezena zákonem. Významnou roli proto hraje iniciativa obecních zastupitelstev a rad, zejména pak osobní angažovanost městských manažerů (starostů, tajemníků, vedoucích odborů). Z konzultací k této problematice s pracovníky městských úřadů z několika českých měst vyplynulo, že některá města dříve šetření spokojenosti prováděla, ale po odchodu iniciátora v nich již nepokračovala.

Ač jsou tato šetření spokojenosti občanů v podstatě dobrovolná, tak o jejich významu vypovídá např. kontrolní závěr Nejvyššího kontrolního úřadu týkající se prostředků státního rozpočtu vynaložených na Národní plán vyrovnávání příležitostí pro občany se zdravotním postižením: „Při kontrole bylo zjišťováno, zda vybraní příjemci dotace provádějí šetření spokojenosti klientů s poskytovanými službami a sledují tak maximální účelnost jimi vynakládaných prostředků dotovaných v rámci NPVP. Odborně prováděná šetření spokojenosti klientů s poskytovanými službami jako základní (a v řadě případů i jediný) způsob hodnocení účelnosti poskytovaných služeb osobám se zdravotním postižením by podle názoru NKÚ měla být obligatorním postupem u organizací využívajících dotace v rámci NPVP.“ NKÚ (2004)

1.4.3 Současné postupy zpracování šetření spokojenosti

Výsledky šetření spokojenosti občanů jsou obvykle zpracovávány pomocí běžných statistických metod. Každý dotazovaný aspekt spokojenosti je při nich vyhodnocován samostatně. Toto zpracování je buď prezentováno již jako konečný výstup, někdy doplněný komentářem, nebo je podkladem pro následné sociologické analýzy. Některá šetření umožňují i porovnání s výsledky v jiných městech. Mezi obvyklé formy měření spokojenosti patří:

- aritmetický průměr - určí se počet respondentů a vypočte se výsledná známka pro jednotlivé otázky, popř. známka za všechny otázky jako celkové hodnocení;
- četnost zastoupení odpovědí respondentů (poměrově v procentech nebo absolutně);
- pořadí šetřených aspektů spokojenosti podle subjektivního významu pro respondenty;
- kombinace výše uvedených.

Lze tedy říci, že postupy, které v současné době běžně využívá regionální management k měření spokojenosti občanů, jsou z hlediska zpracování jednoduché. Mají své nesporné výhody, např. jsou nenáročné z hlediska použitých informačních technologií, potřebné technologie zvládnou vlastní pracovníci a závěry jsou dobře srozumitelné a sdělitelné.

Jejich nedostatkem je potom nedostatečné využití možností, které získaná data nabízejí. Při šetřeních jsou shromažďovány velké objemy dat, která v sobě ukrývají významný informační potenciál[5] a nad kterými je možné formulovat různé typy úloh. Nabízí se zde tedy široké pole působnosti pro jejich využití pomocí umělé a výpočetní inteligence.

[5] Překážkou využití shromážděných dat je často jejich nedostupnost, pokud jsou data majetkem subjektu, který šetření prováděl. Obvykle se zveřejňují jen výsledky výzkumů či šetření, nikoliv „surová“ data.

2 Návrh úloh zpracování dat o spokojenosti občanů

Vzhledem k povaze problematiky se jeví jako výhodné přistoupit ke zpracování dat o spokojenosti občanů jako k úlohám z oblasti dobývání znalostí z databází - data miningu. Data mining definuje, podle HEARST (2003), Usama Fayyad jako „netriviální proces zjišťování platných, neznámých, potenciálně užitečných a snadno pochopitelných závislostí v datech“. Přes relativní mládí (začátek 90. let minulého století) má tato disciplína vypracované metodiky, BERKA (2003), které si kladou za cíl poskytnout uživatelům jednotný rámec pro řešení různých úloh dobývání znalostí. Jde především o metodiky 5A (Assess, Access, Analyze, Akt, Automate), SEMMA (Sample, Explore, Modify, Model, Assess) a CRISP-DM (Cross Industry Standard Process for Data Mining).

Pro další postup byla zvolena metodika CRISP-DM, podle které lze postupovat bez ohledu na software použitý pro další zpracování. CRISP-DM člení celý proces data miningu do šesti základních etap. Jsou to, CHAPMAN (2000), porozumění problematice, porozumění datům, příprava dat, modelování, hodnocení a využití výsledků.

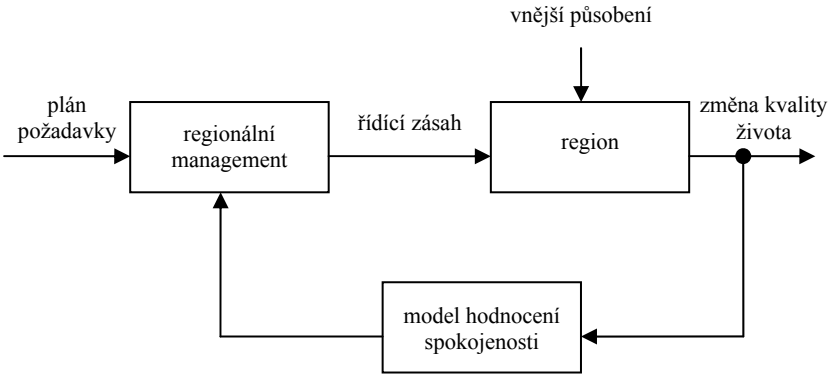
2.1 Iniciativa Společné evropské indikátory

Nejrozsáhlejším projektem na území České republiky, který se týká měření spokojenosti obyvatel, je projekt občanského sdružení TIMUR - Týmová iniciativa pro místní udržitelný rozvoj. Cílem TIMUR je podporovat udržitelný rozvoj měst, obcí a jejich sdružení v ČR za pomoci zavádění místních indikátorů udržitelného rozvoje. Partnerem TIMUR se může stát město (obec, mikroregion, výhledově kraj), které má zájem provádět aktivní politiku v oblasti udržitelného rozvoje, zapojovat veřejnost do rozhodování o rozvoji města a sleduje indikátory udržitelného rozvoje. V roce 2003 byl zahájen pilotní projekt zapojením měst Hradce Králové a Vsetína. Velkým přínosem této iniciativy je koordinace šetření spokojenosti občanů, poskytnutí jednotné metodiky a zpracování výsledků dotazníkových průzkumů. Kromě zjištění trendu spokojenosti za určité období tak lze provádět i porovnání (benchmarking) se srovnatelnými městy či regiony.

Vhodným podkladem pro zpracování dat s využitím umělé a výpočetní inteligence jsou data získaná dotazníkovým průzkumem Indikátor A1 - Spokojenost s místním společenstvím. Tento indikátor patří v rámci „Indikátorů udržitelného rozvoje na místní úrovni“

do skupiny indikátorů „Společné evropské indikátory“. Společné evropské indikátory jsou sada deseti indikátorů, které odráží rozličné aspekty života a řízení města (základ je ve třech pilířích udržitelného rozvoje[6]). Umožňuje sběr srovnatelných údajů v rámci celé Evropy a v rámci srovnatelně velkých sídelních útvarů. Projekt vychází z iniciativy skupin kolem Evropské komise. NOVÁK (2006), PRACOVNÍ SKUPINA (2000).

První fází CRISP-DM, jak uvádí BERKA (2003) a CHAPMAN (2000), je pochopení cílů úlohy z manažerského hlediska a její převod na úlohu dobývání znalostí. V tomto případě je v manažerské roli regionální management. Manažerskou úlohou je hledání souvislostí mezi demografickými údaji o občanech, dílčími aspekty spokojenosti občanů s místním společenstvím, aspekty spokojenosti občanů s různými rysy samosprávy a spokojeností celkovou. Pokud by se podařilo definovat skupiny občanů, kteří jsou nespokojeni, mohou získané znalosti sloužit jako podklady pro stanovení priorit při dalším rozhodování. Model hodnocení spokojenosti občanů jako subsystém systému řízení je znázorněn na obr. 11.



Obr. 11 Model hodnocení spokojenosti občanů jako subsystém systému řízení

Z hlediska dobývání znalostí lze úlohu formulovat např. jako:

- nalezení takových druhů demografických údajů o občanech, které umožní určit míru jejich celkové spokojenosti;
- nalezení takových aspektů spokojenosti, které umožní určit míru jiného (cílového) aspektu spokojenosti;

[6] Udržitelný rozvoj je novým rámcem strategie civilizačního rozvoje. Vychází z klasické a široce přijaté definice Komise OSN pro životní prostředí a rozvoj z r. 1987, která považuje rozvoj za udržitelný tehdy, naplní-li potřeby současné generace, aniž by ohrozil možnosti naplnit potřeby generací příštích. Cílem je takový rozvoj, který zajistí rovnováhu mezi třemi základními pilíři: sociálním, ekonomickým a environmentálním, jak symbolicky vyjádřilo jeho heslo: lidé, planeta, prosperita. SUR (2004)

- nalezení takových aspektů spokojenosti občanů, které nejvíce přispívají k celkové spokojenosti.

Úloha proto může být pojata jako klasifikační - cílem je zařazení občanů do tříd podle míry jejich spokojenosti.

2.2 Dotazníkový průzkum Spokojenost občanů s místním společenstvím v Chrudimi

Město Chrudim již delší dobu využívá výsledků dotazníkových průzkumů ze sady Společné evropské indikátory jako zpětné vazby při řízení a rozvoji města. Dávají se na vědomí radě a zastupitelstvu, odborům městského úřadu, organizacím města (Technické služby) a jsou zveřejňovány. Výsledky jsou využívány např. při rozhodování o přidělování dotací nebo jako argument při jednáních o obchvatu města.

Pro vybrané úlohy lze využít indikátor A1 - Spokojenost občanů s místním společenstvím (Obecná spokojenost občanů s různými rysy samosprávy), který zjišťuje a vyčísluje subjektivní pocit spokojenosti občanů s městem, ve kterém žijí a pracují, a dílčí aspekty této spokojenosti. Dotazníkový průzkum zaměřený na tento indikátor byl proveden na území města Chrudim na podzim roku 2007. Předchozí průzkum tohoto indikátoru byl proveden v roce 2004. Záznamy z papírových dotazníků (viz Příloha 1: Dotazník – Indikátor 1) jsou prvotně ukládány v aplikaci MS Access, prezentovány jsou pak jako jednoduché statistické přehledy ve formě tabulek a grafů zpracovaných v MS Excel.

Sociologický empirický výzkum, GIDDENS (1999), VESELÁ (2002), získává konkrétními metodami a technikami empirická data v sociální a sociálně psychické realitě a zpracovává je. Poskytuje podklady jak pro teoretickou práci, tak pro rozhodování a řízení této reality. Právě zaměření na konkrétní aktivity k dosahování žádoucích změn ve zkoumané realitě je pro empirický výzkum příznačné. Technika sociologického výzkumu je zpravidla připravena tak, že představuje více či méně dokonalý systém znaků, pomocí nichž jsou zjišťovány a měřeny vlastnosti sociálního objektu. Metodou pro získání dat, která jsou použita v této práci, je výzkum dotazníkem. Specifičnost této metody spočívá v tom, že předmětem analýzy se stávají písemné odpovědi na pevné a pro všechny dotazované stejně formulované otázky.

Sociologie využívá dotazníková šetření k přehledovým studiím. Technika dotazníku má své výhody - umožňuje efektivní shromáždění dat od velkého množství osob, umožňuje

jednoznačnou formulaci otázek, přesná srovnání odpovědí a hromadné zpracování. Nevýhodou však je, že pokud je dotazník vysoce standardizován, mohou být důležité rozdíly ve stanoviscích respondentů chybně interpretovány, při eventuálním neporozumění nelze obvykle už nic opravit, doplnit nebo zpřesnit. Odpovědi mohou také odrážet to, co o sobě lidé tvrdí, a ne to, co si skutečně myslí.

Některé tyto skutečnosti jsou zřejmé i při bližším zkoumání odpovědí v dotaznících indikátoru A1 a je třeba jim věnovat pozornost při předzpracování dat.

3 Metody využitelné pro řešení úloh o spokojenosti občanů

Data získaná při šetření spokojenosti občanů lze zpracovávat - podle formulace úlohy - pomocí širokého spektra metod umělé a výpočetní inteligence. Kapitola je proto zaměřena jen na oblasti, které jsou zahrnuty v aplikační části práce a které užívají klasifikační systémy založené na pravidlech. Jsou to rozhodovací (klasifikační) stromy[7] a fuzzy inferenční systém (FIS) typu Mamdani. Algoritmy pro vytváření rozhodovacích stromů jsou v rámci umělé inteligence řazeny do oblasti rozpoznávání obrazů a strojového učení. Fuzzy inferenční systémy patří v rámci výpočetní inteligence do oblasti fuzzy množin. BERKA (2003), OLEJ (2003).

3.1 Pravidlové systémy

Znalosti v pravidlových systémech, jinak též systémech založených na pravidlech, jsou reprezentovány pomocí pravidel typu JESTLIŽE - PAK (IF - THEN), která mohou mít takovéto tvary:

IF předpoklad THEN závěr,

IF situace THEN akce,

IF podmínka THEN závěr AND akce,

IF podmínka THEN důsledek1 ELSE důsledek2,

IF podmínka1 AND podmínka2 THEN závěr,

IF podmínka1 OR podmínka2 THEN závěr,

IF předpoklad THEN závěr WITH váha.

Část pravidla za IF (levá strana pravidla) se nazývá antecedent, podmínková část nebo také část vzorů. Tato část může sestávat z individuálních podmínek, resp. vzorů. Část pravidla za THEN (pravá strana pravidla) se nazývá konsekvent a může také obsahovat několik akcí nebo závěrů. Váha je subjektivní měrou důvěry experta v platnost daného pravidla.

[7] V české terminologii se pod pojmem rozhodovací stromy skrývá i úplně jiný nástroj - rozhodovací stromy jsou také nástrojem rozhodovací analýzy (Decision Tree Analysis). Jsou vhodné pro zobrazení a podporu řešení víceetapových rozhodovacích procesů za rizika, resp. nejistoty. V některých pramenech lze nalézt termín klasifikační strom, což je pro rozhodovací stromy užívané ke klasifikaci přesnější výraz. FOTR (2006).

V předpokladové části se mohou vyskytovat spojky AND a OR, v důsledkové části mohou být spojky AND a ELSE. Součástí pravidla může být také tzv. kontext, ve kterém má být pravidlo uvažováno. DVOŘÁK (2004), MAŘÍK – ZDRÁHAL (1997).

Tato pravidla se nazývají **podmíněná pravidla**. Užívá se i názvů produkční, odvozovací, rozhodovací, reaktivní, asociační, if-then pravidlo. Podmíněné pravidlo je výsledkem logické operace implikace. Implikace, KLIMEŠ (1985), tvoří složený výrok ze dvou výroků pomocí logické spojky, která odpovídá spojce „jestliže - pak“ a značí se symbolem „$\Rightarrow$“. Je to logický vztah mezi výroky, z nichž jeden je konsekvent (důsledek) druhého neboli antecedentu (předpokladu). Antecedent (uváděný slovem jestliže) stanovuje, za jakých podmínek konsekvent platí. Podmíněná pravidla se v obecném tvaru zapisují

$$E \Rightarrow H, \tag{3.1}$$

kde E představuje pozorování (evidence) a H znamená hypotézu (hypothesis), nebo

$$Ant \Rightarrow Suc, \tag{3.2}$$

kde Ant představuje předpoklad a Suc závěr.

Odlišnosti najdeme u asociačních pravidel. Asociační pravidla, BERKA (2003), SHEN (2006), ve tvaru $X \Rightarrow Y$ jsou výsledkem implikace představující podmínku současného výskytu položek množiny X a Y v dané databázi. Hledáme kombinace

$$Ant \wedge Suc, \tag{3.3}$$

přičemž není žádná položka množiny upřednostňována jako závěr pravidla

3.2 Klasifikační systémy založené na pravidlech

Úloha **rozpoznávání**, KOTEK(1993), spočívá v zařazování objektů reálného světa do tříd. Nejprve je třeba stanovit hledisko, ze kterého budou objekty posuzovány, tzn. určit veličiny, které je charakterizují. Definici systému provádí řešitel úlohy na základě analýzy vlastností objektů a tříd. Data naměřená na objektu uspořádáme do vektoru, který nazveme obraz. Ten „zobrazuje“ objekt z hlediska zvolené definice systému. Prostor všech obrazů nazveme obrazový prostor.

Úlohu rozpoznávání lze rozdělit na dvě části:

- zpracování dat naměřených na objektu (zpracování obrazu) tak, aby byla maximalizována diskriminační schopnost při minimalizaci objemu dat;

- přiřazení indikátoru třídy jednotlivým popisům získaným v bodě prvním, tj. vlastní klasifikace.

Obě části úlohy spolu těsně souvisí, výsledkem zpracování dat musí být reprezentace vyhovující zvolené klasifikační metodě. Příznakové metody odvozují z obrazu kvantitativní ocenění (příznaky) podstatných vlastností objektů. Strukturální metody reprezentují objekt pomocí popisných elementů (primitiv - nejjednodušších kvalitativních charakteristik) a vztahů mezi nimi (např. relace „být před“). KOTEK(1993), MAŘÍK (1993)

Klasifikační úlohy, ŽELEZNÝ - KLÉMA - ŠTĚPÁNKOVÁ (2003), jsou typem úloh prediktivních. V nich uživatele zajímá hodnota cílové veličiny v závislosti na jiných, známých veličinách (atributech). Prediktivní modely mají za cíl předpovídat neznámou hodnotu na základě známých hodnot charakterizujících objekt predikce. Pro každou konkrétní úlohu je množina uvažovaných veličin jednoznačně rozdělena na veličinu závislou (výstup) a veličiny nezávislé (vstupy, atributy). V případě kategoriální výstupní veličiny hledáme funkci příslušnosti ke konceptu a řešený problém označujeme jako klasifikační úlohu.

Klasifikační systém založený na pravidlech je klasifikační systém, ve kterém jsou znalosti vyjádřeny ve tvaru podmíněných rozhodovacích pravidel. Syntaxe pravidla je tedy, BERKA (2003),

$$IF\ předpoklad\ THEN\ klasifikace \tag{3.4}$$

nebo

$$Ant \Rightarrow Class, \tag{3.5}$$

kde *Ant* (předpoklad) je kombinace vytvořená z kategorií vstupních atributů a *Class* (závěr) je zařazení příkladu do třídy, tedy nějaká kategorie cílového atributu. Rozhodovací pravidla jsou pro klasifikaci využívána v následujících skupinách metod:

- algoritmy pro vytváření rozhodovacích (klasifikačních) stromů,
- algoritmy pro vytváření rozhodovacích pravidel,
- fuzzy inferenční systémy,
- hrubé množiny (Rough Sets).

První tři jsou popsány v části 3.4 a části 3.5. Hrubé množiny[8] jsou matematický nástroj pro analýzu vágních a nepřesných dat. Základním poznatkem je, že vágnost nebo nepřesnost vede k nerozlišitelnosti, která formálně reprezentuje nedostatek znalostí, a je základní myšlenkou

[8] Hrubé množiny představil Zdzdislawa Pawlak v práci Rough Sets v roce 1982.

podtrhující filozofii hrubých množin. Relace nerozlišitelnosti se používá k definování základních operací na množinách – dolní a horní aproximaci množiny, které se využívají místo přesných pojmů. S každou množinou dat pak nakládáme jako se speciálním druhem rozhodovací tabulky obsahující rozhodovací pravidla. Dolní a horní aproximace se používají k analýze vlastností rozhodovacích tabulek. Při vytváření modelů není třeba žádná další informace o datech, jako např. distribuční funkce nebo funkce příslušnosti. GRZYMALA-BUSSE - ZIARKO (2006), NOVOTNÝ (2006)

Klasifikační systémy, které nejsou založeny na pravidlech, využívají metody jako jsou neuronové sítě, genetické algoritmy, diskriminační analýza, bayesovské klasifikace a metody založené na analogii, např. případové usuzování (Case-Based Reasoning, CBR), pravidlo nejbližšího souseda, shlukování a další.

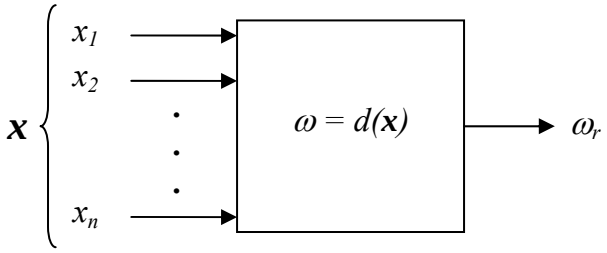
Matematický model objektu odvozený od souboru měřených veličin, obraz ***x*,** je reprezentován sloupcovým vektorem

$$\boldsymbol{x} = \left[x_1, x_2, \ldots, x_n\right]^T \tag{3.6}$$

jehož složky $x_1, x_2, \ldots, x_n$ se nazývají příznaky a odpovídají jednotlivým měřeným veličinám. Množině výstupů odpovídá množina symbolů ω_r, které se jmenují indikátor třídy.

Klasifikátor je stroj, který provádí klasifikaci - zobrazuje příznakový prostor na množinu indikátorů tříd (obr. 12). Pojem „stroj“ chápeme v širším slova smyslu - označujeme jím i algoritmus atd. Klasifikátor lze nastavit dvěma způsoby, KOTEK (1980, 1993):

- analýzou problému a definováním rozhodovacího pravidla před klasifikací;
- vytvořením rozhodovacího pravidla s použitím objektů, jejichž správná klasifikace je předem známá, tj. učení s učitelem (např. algoritmy pro vytváření rozhodovacích stromů, FIS typu Mamdani).



Obr. 12 Klasifikátor, KOTEK (1980)

Vztah mezi vstupem a výstupem klasifikátoru, KOTEK (1980), popisuje skalární funkce vektorového argumentu

$$\omega = d(\mathbf{x}). \tag{3.7}$$

Tato funkce se nazývá rozhodovací pravidlo. Obrazový prostor se rozhodovacím pravidlem rozděluje na R disjunktních podmnožin $\mathcal{R}_r$, $r = 1, ..., R$, přičemž podmnožina $\mathcal{R}_r$ obsahuje ty obrazy $\mathbf{x}$, pro které $\omega_r = d(\mathbf{x})$. Nadplochy, které od sebe oddělují sousední množiny $\mathcal{R}_r$ a $\mathcal{R}_s$ se nazývají rozdělující nadplochy. Určení rozdělujících nadploch se tak stává základním problémem návrhu klasifikátoru.

3.3 Formalizace klasifikační úlohy

Kvalita rozhodování je vždy spojena s kvalitou a množstvím informace, kterou má rozhodující orgán, ať člověk nebo stroj, k dispozici. Jinak řečeno, pravděpodobnost chyby klasifikátoru je nepřímo úměrná množství informace, obsažené ve vstupních obrazech. Proto by měl být obraz co nejúplnější, ale počet příznaků nelze zvyšovat, z více důvodů, do nekonečna. Vzrůstající počet příznaků navíc nezaručuje zvyšování správnosti klasifikace. Zcela na místě je proto řešení otázky volby informativních příznaků, které se výrazně podílejí na správnosti klasifikace. Volba informativních příznaků není dosud formalizována. Neexistuje teorie, která by nám předem dovolila stanovit, které veličiny na daných předmětech měřit. MAŘÍK (1993), KOTEK (1993)

Současné teorie nabízejí dvě řešení:

- výběr minimálního množství příznaků z předem zvolené množiny příznaků (velký vliv zde má zkušenost a intuice konstruktéra - např. FIS typu Mamdani);
- uspořádání příznaků v předem zvolené množině příznaků do posloupnosti podle množství informace, které jednotlivé příznaky nesou (např. algoritmy pro vytváření rozhodovacích stromů).

Je zřejmé, že terminologie klasifikačních úloh není ustálená a vyplývá často z konkrétních pojmů užívaných ve zkoumaném oboru. Nadále jsou užívány tyto pojmy:

- **objekt,** pro měřené objekty reálného světa (jinak též záznam databáze, příklad, případ, předmět);
- **atribut**, pro veličiny, které objekty charakterizují (proměnná, znak, příznak, aspekt, vlastnost).

Analyzovaná data , BERKA (2003), KOTEK (1980), jsou uložena v matici $\boldsymbol{D}$, tvořené n řádky a m sloupci

$$D = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{11} & \dots & x_{nm} \end{bmatrix}. \tag{3.8}$$

Řádky matice reprezentují sledované objekty, i-tý objekt je tedy řádek $\boldsymbol{x}_i$

$$\boldsymbol{x_i} = [x_{i1}, x_{i2}, \dots, x_{im}]. \tag{3.9}$$

Sloupce datové matice odpovídají atributům, j-tý atribut (j-tý sloupec) je značen symbolem $\boldsymbol{A}_j$

$$\boldsymbol{A}_j = [x_{1j}, x_{2j}, \dots, x_{nj}]^T. \tag{3.10}$$

Jak již bylo řečeno, u klasifikačních úloh předpokládáme, že existuje atribut obsahující informaci o zařazení objektů do tříd - cílový atribut $\boldsymbol{C}$ (jinak též závislá proměnná, vysvětlovaná veličina)

$$\boldsymbol{C}: [y_1, y_2, \dots, y_n]^T. \tag{3.11}$$

Ostatní, necílové atributy $\boldsymbol{A}_j$ nazýváme vstupní atributy (jinak též nezávislá proměnná, vysvětlující veličina).

Popis konceptu, BERKA (2003), který byl nalezen na základě použitých příkladů, nemusí odpovídat jiným (dosud nezpracovaným) příkladům téhož konceptu. Proto se obvykle data použitá při získávání znalostí rozdělují na data trénovací a data testovací. Trénovací data se použijí ve fázi učení, testovací pak slouží k prověření získaných znalostí. Přidáme-li cílový atribut do datové matice $\boldsymbol{D}$, získáme data vhodná pro použití metod učení s učitelem - trénovací data, značená $\boldsymbol{D}_{TR}$

$$\boldsymbol{D}_{TR} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} & y_1 \\ x_{21} & x_{22} & \dots & x_{2m} & y_2 \\ \vdots & \vdots & & \vdots & \vdots \\ x_{n1} & x_{11} & \dots & x_{nm} & y_n \end{bmatrix}. \tag{3.12}$$

Objekt (trénovací příklad) z této matice je značen

$$\boldsymbol{o}_i = [\boldsymbol{x}_i, y_i]. \tag{3.13}$$

Data jsou považována za množinu objektů

$$\boldsymbol{D}_{TR} = \{\boldsymbol{o}_i, i = 1, \dots, n\}. \tag{3.14}$$

Předpokládáme, že pro každý objekt $\boldsymbol{o}_i$ známe všechny hodnoty $\boldsymbol{x}_i$ i hodnotu y_i. Ve většině případů nezáleží na pořadí objektů v datové matici. Výjimku tvoří prostorová nebo časová data, kdy uspořádání vyplývá z povahy těchto dat.

3.4 Rozhodovací stromy

K nejznámějším klasifikačním modelům patří algoritmy pro vytváření rozhodovacích stromů. Rozhodovací strom, RYCHLÝ (2006), můžeme definovat jako strom (stromový graf), kde každý nelistový uzel stromu představuje test hodnoty atributu a větve vedoucí z tohoto uzlu možné výsledky testu. Listové uzly stromu jsou ohodnoceny identifikátory tříd (výsledky klasifikace). Vlastní klasifikace pomocí stromu probíhá cestou záznamu od kořene stromu k jeho listu. V každém kroku je záznam otestován podle testu v aktuálním uzlu rozhodovacího stromu a dále pokračuje po větvi shodné s konkrétním výsledkem testu. Pokud takto záznam dojde až do listového uzlu, je oklasifikován třídou identifikovanou hodnotou příslušného listu rozhodovacího stromu.

Při tvorbě rozhodovacího stromu[9] se tedy postupuje metodou „rozděl a panuj“. Trénovací data se postupně rozdělují na menší a menší podmnožiny tak, aby v těchto podmnožinách převládaly příklady jedné třídy. Atribut vhodný pro větvení stromu vybíráme na základě jeho charakteristik převzatých z teorie informace a pravděpodobnosti: entropie, informačního zisku, poměrného informačního zisku, χ^2 - testu, Giniho indexu a dalších. Intuitivní vizuální zobrazení stromem pomáhá jasnějšímu pochopení výsledků a vztahů i laickým uživatelům a v praxi tak usnadňuje jejich rozhodování. Stromové grafy dovolují vizuálně prozkoumat výsledky a posoudit vhodnost modelu. Rozhodovací strom lze poměrně snadno převést na rozhodovací pravidla. Každé cestě stromem od kořene k listu odpovídá jedno pravidlo. Nelistové uzly jsou předpoklady, listový uzel pak závěrem pravidla. Existují i samostatné algoritmy pro přímou tvorbu pravidel - bez využití rozhodovacího stromu. Jedním z nejznámějších algoritmů pro tvorbu pravidel je algoritmus pokrývání množin, známý jako AQ[10], pracující metodou „odděl a panuj“. Při pokrývání množin jde totiž o to nalézt pravidla, která pokrývají příklady téže třídy a oddělit je od příkladů třídy jiné. BERKA (2003), SPSS (2008).

[9] Autorem obecného algoritmu TDIDT a algoritmu ID3 je John Ross Quinlan.

[10] Autorem je Ryzsard Michalski

Mezi nejznámější a běžně užívané algoritmy pro vytváření rozhodovacích stromů, BERKA (2003), SPSS (2008), patří:

- C5.0, nejnovější verze řady algoritmů jako ID3 a C4.5 , které v zásadě pracují podle algoritmu TDIDT (Top Down Induction of Decision Trees); díky zpřístupnění zdrojových kódů programu se C4.5 stal de facto standardem pro tvorbu rozhodovacích stromů;
- CHAID (Chi-square Automatic Interaction Detection), statistický algoritmus založený na optimální hodnotě χ^2 - testu závislosti nebo F-testu, štěpí skupiny vždy na vhodný počet statisticky homogenních podskupin;
- Exhaustive CHAID, modifikace algoritmu CHAID, která prozkoumává všechna možná štěpení pro každý prediktor;
- C&RT (Classification & Regression Trees), vytváří binární stromy (skupina je štěpena vždy na dvě části) a maximálně homogenní skupiny podle statistických kriterií;
- QUEST - statistický algoritmus, který vybírá proměnné nevychýleně a rychle, efektivně vytváří přesné binární stromy.

3.5 Fuzzy inferenční systémy

V posledních letech, ZADEH (1998), jsme svědky rychlého růstu počtu a různorodosti aplikací fuzzy logiky od spotřebního zboží k řízení průmyslových procesů a systémů na podporu rozhodování. Pojem fuzzy logika má dva různé významy. V užším smyslu je fuzzy logika logický systém, který je rozšířením vícehodnotové logiky. V širším smyslu, dnes převládajícím, je fuzzy logika téměř synonymem k teorii fuzzy množin[11]. Hlavním zdrojem úspěchu, NOVÁK (2000), je možnost zahrnout nepřesnost a poměrně jednoduchým způsobem pracovat s významy slov přirozeného jazyka.

[11] Anglický termín fuzzy znamená „neurčitý“, „neostrý“, „rozmazaný“ a nepřekládá se.

3.5.1 Fuzzy množiny a fuzzy usuzování

Při definici fuzzy množiny, NOVÁK (2000), nejprve definujeme množinu U, nazývanou univerzum, což může být množina libovolných prvků. Fuzzy množina je funkce

$$A: U \rightarrow [0, 1]. \tag{3.15}$$

Fuzzy množina je tedy tvořena prvky x vybíranými z množiny U, $x \in U$, z nichž každý má přiřazeno číslo $a \in [0, 1]$ nazývané stupeň příslušnosti prvku x do fuzzy množiny A. Zároveň je $A(x)$ stupeň pravdivosti toho, že x patří do A. Vztah (3.15) se někdy nazývá funkce příslušnosti a fuzzy množina je tedy ztotožněna se svou funkcí příslušnosti. Křivka funkce příslušnosti může mít různý tvar, užívá se tvar trojúhelníkový, lichoběžníkový, zvonový, tvar Gaussovy křivky nebo sigmoidální funkce. Tvary křivek jsou podrobně definovány v knihách JURA (2003), NOVÁK (2000), OLEJ (2003), ROSS (2004). Zvláštním tvarem funkce příslušnosti je singleton, JURA (2003), ROSS (2004), který se využívá pouze pro výstupní funkce příslušnosti. Funkce příslušnosti lze značit také μ, OLEJ (2003), ROSS (2004), potom se stupeň příslušnosti prvku x do fuzzy množiny A zapisuje jako $\mu_A(x)$.

Explicitně se fuzzy množiny zapisují

$$A = \{a_1/x_1, \ldots, a_n/x_n\}. \tag{3.16}$$

Není-li univerzum konečná množina, zapisujeme fuzzy množinu

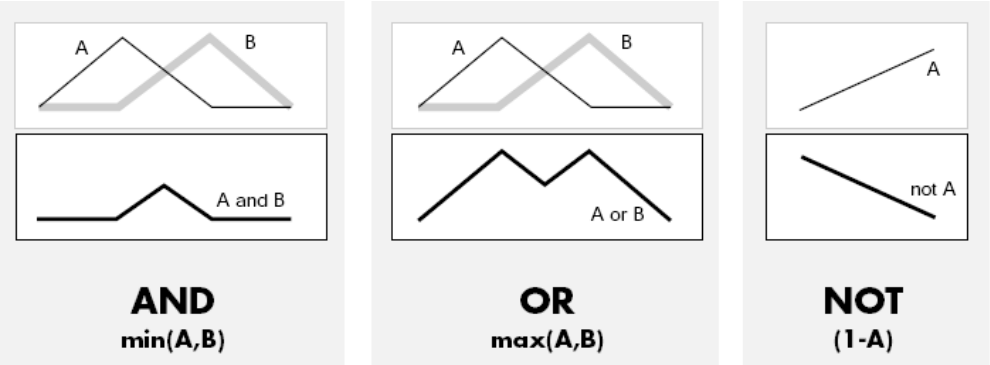
$$A = \{a_i/x_i \mid i \in I\}, \tag{3.17}$$

kde I je nějaká indexová množina, popř. jsou-li prvky x reálná čísla a stupně příslušnosti jsou dány nějakou funkcí, lze zapsat fuzzy množinu

$$A = \{f(x)/x \mid x \in R\}, \tag{3.18}$$

kde R je množina všech reálných čísel.

Pro fuzzy množiny lze definovat, mimo dalších operací, základní operace průniku nebo konjunkce (AND), sjednocení nebo disjunkce (OR) a doplňku (NOT). Tyto operace jsou pro fuzzy množiny definovány více způsoby, jeden z možných, pro fuzzy množiny A a B, je uveden na obr. 13.



Obr. 13 Operace mezi fuzzy množinami, THE MATHWORKS (1998)

Fuzzy pravidla typu IF - THEN mají aplikace prakticky ve všech oblastech od řízení po rozpoznávání obrazů a predikci vývoje ekonomiky. Tato pravidla, JURA (2003), představují úsudek (odvození, inference), jehož vstupní informace jsou vágní. Mají tvar

$$IF\ (fuzzy\ výrok)\ THEN\ (fuzzy\ výrok), \tag{3.19}$$

např.

$$IF\ (věk\ je\ střední)\ THEN\ (spokojenost\ je\ vysoká), \tag{3.20}$$

a představují implikaci, nikoliv ostrou, ale fuzzy implikaci. Výroky předpokladové části pravidla mohou být propojeny logickými spojkami. Výsledné rozhodnutí potom představuje agregaci výsledků všech pravidel.

3.5.2 Struktura fuzzy inferenčních systémů

Fuzzy systém, JURA (2003), je takový systém, jehož proměnné nabývají hodnot, které nejsou definovány ostrými čísly, nýbrž jsou definovány slovními hodnotami, tj. fuzzy množinami. Jednoduchý základní tvar fuzzy pravidlového systému, který uvádí ROSS (2004), je popsán tabulkou 1. Základní jednotkou reprezentující ve fuzzy systémech znalost je jazyková proměnná. To je taková proměnná, jejíž hodnoty jsou slova nebo věty přirozeného jazyka.

Tab. 1 Základní tvar fuzzy pravidlového systému, upraveno podle ROSS (2004)

pravidlo 1:	IF podmínka C^1, THEN omezení R^1
pravidlo 2:	IF podmínka C^2, THEN omezení R^2
.	
.	
.	
pravidlo *r*:	IF podmínka C^r, THEN omezení R^r

Vstupní i výstupní proměnné fuzzy odvozování, JURA (2003), jsou jazykové proměnné, jejichž hodnoty jsou vyjádřeny pomocí fuzzy množin. Na základě vyhodnocení všech fuzzy pravidel je nalezen výsledek přibližného úsudku - inference. Proto se tyto systémy nazývají fuzzy inferenční systémy (FIS). V některých aplikacích získáváme vstupní hodnoty ve formě ostrých čísel a také výstupní hodnoty jsou požadovány ve formě ostrých čísel. V tomto případě je třeba doplnit fuzzy systém o modul fuzzifikace a modul defuzzifikace. Cílem fuzzifikace je transformovat ostrou vstupní hodnotu proměnné x na fuzzy množinu. Cílem defuzzifikace je získat z fuzzy množiny, která je výsledkem inferenčního mechanizmu, ostrou výstupní hodnotu. Procesy jsou definovány v knihách JURA (2003), NOVÁK (2000), OLEJ (2003), ROSS (2004). Všeobecná struktura fuzzy inferenčního systému potom obsahuje, OLEJ (2003), proces fuzzifikace, návrh báze podmíněných pravidel nebo jejich automatickou extrakci ze vstupních údajů, aplikaci operátorů v podmíněných pravidlech, implikaci a agregaci v rámci těchto pravidel a proces defuzzifikace.

Mezi tři běžné inferenční metody patří systémy Mamdani, Sugeno a Tsukamoto. ROSS (2004).

Pravidlo v systému Mamdani, který má dva vstupy x_1 a x_2 a jeden výstup y, má tvar

$$IF\ x_1\ je\ A_1\ AND\ x_2\ je\ A_2\ THEN\ y\ je\ B \tag{3.21}$$

kde A_1 a A_2 jsou fuzzy množiny představující antecedent a B je fuzzy množina představující konsekvent. Výstupem FIS Mamdani je fuzzy množina.

Pravidlo v systému Sugeno (nebo TSK - Takagi, Sugeno a Kang), který má dva vstupy x a y a jeden výstup z, má tvar

$$IF\ x\ je\ A\ AND\ y\ je\ B\ THEN\ z\ je\ z = f(x, y) \tag{3.22}$$

Výstupem FIS Sugeno je ostré číslo.

Třetí metodou je FIS Tsukamoto. V této metodě je konsekvent každého pravidla reprezentován fuzzy množinou s monotónní funkcí příslušnosti.

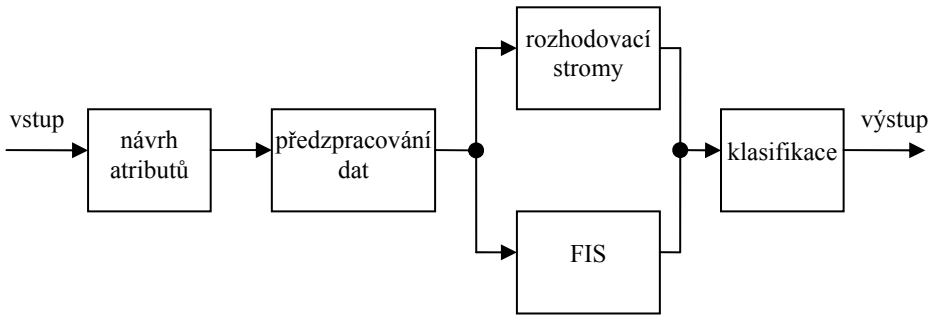
Naplnění báze pravidel je klíčovým momentem celého procesu. Mezi důležité vlastnosti souboru pravidel, JURA (2003), patří úplnost (v souboru nechybí žádné důležité pravidlo) a konzistence (neexistují pravidla, která mají stejný antecedent, ale různý konsekvent).

3.6 Návrh metod řešení klasifikační úlohy o spokojenosti občanů

První vybranou metodou pro řešení klasifikační úlohy o spokojenosti občanů jsou algoritmy pro vytváření rozhodovacích stromů C5.0, CHAID, C&RT a QUEST. Algoritmy jsou popsány v části 3.4. Jako prostředí pro modelování byl zvolen software SPSS Clementine společnosti SPSS. Tento nástroj zahrnuje vybrané algoritmy a podporuje použitou metodiku CRISP-DM.

Spokojenost občanů je rozsáhlý a neurčitý pojem, jako např. vzdělání nebo zdraví, ale obojí je třeba nějak měřit a hodnotit. Je jasné, že někdo je spokojen více než druhý. Působení na zlepšování kvality života občanů a tím i zvyšování míry jejich spokojenosti vyžaduje, abychom věděli, podle čeho lze usoudit na úspěch tohoto působení. Vzhledem k nepřesnosti pojmů užívaných při měření spokojenosti občanů je vhodné využití fuzzy inferenčního systému Mamdani pro jeho schopnost tuto nepřesnost zahrnout. FIS typu Mamdani je proto druhou vybranou metodou. FIS je popsán v části 3.5. Jako prostředí pro modelování byl zvolen simulační software MATLAB společnosti The MathWorks. Jeho součástí je Fuzzy Logic Toolbox, který je nástrojem pro práci s fuzzy množinami.

Jsou-li navrženy metody řešení klasifikační úlohy, je možné navrhnout model hodnocení spokojenosti občanů. Ten je zahrnut jako subsystém v systému řízení a znázorněn na obr. 11 v části 2.1. Na nižší rozlišovací úrovni je tento subsystém znázorněn na následujícím obr. 14.



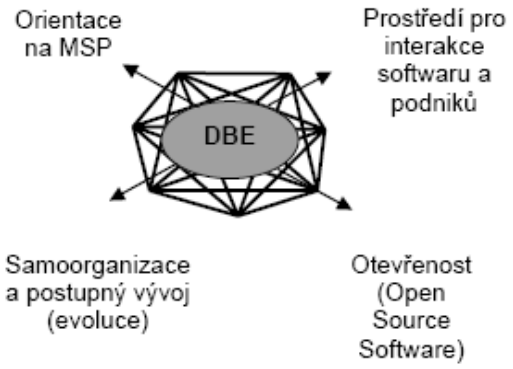
Obr. 14 Model hodnocení spokojenosti občanů

4 Využití Digitálního ekosystému pro podnikání v oblasti veřejné správy

Digitální ekosystém pro podnikání (The Digital Business Ecosystem – DBE), ERISA (2007), je internetové softwarové prostředí s volným přístupem a s otevřeným zdrojovým kódem (Open Source), ve kterém mohou být vyvíjeny a provozovány podnikové aplikace. Vytváří infrastrukturu, ve které všechny firmy (zejména malé a střední podniky - MSP) mohou spolupracovat a soutěžit. Umožní jim to:

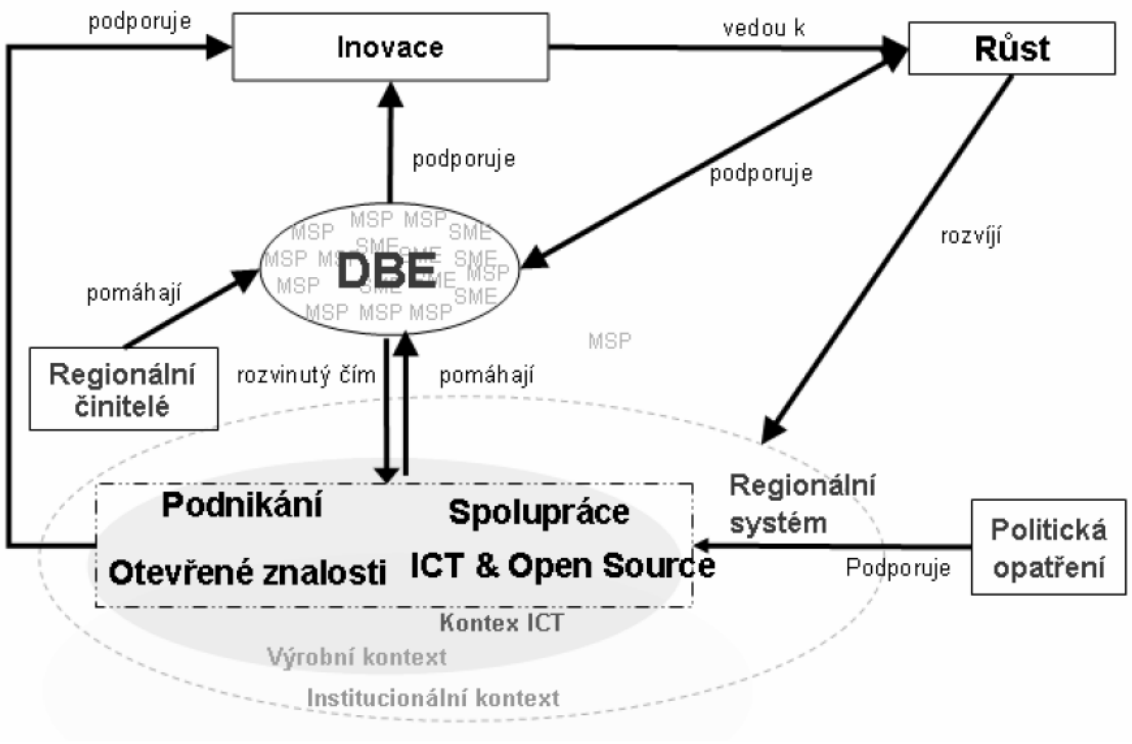
- snadný vývoj služeb a jejich napojování na tzv. „legacy“ aplikace (tj. „zděděné“ aplikace, které uživatelé nemohou opustit, protože jsou pro chod firmy nepostradatelné);
- zpřístupnění (zveřejnění) služeb a aplikací, které dosáhly jednoduché formální specifikace (např. používají strukturovaný přirozený jazyk – angličtinu) a obsahují jakoukoliv požadovanou relevantní informaci (jako jsou např. používané obchodní modely, licenční podmínky), protože jak proprietární, tak zdrojově otevřené aplikace koexistují;
- využití služeb a aplikací v kombinaci, protože infrastruktura DBE podporuje skládání, spojování a migraci služeb a znalostí.

DBE tedy pomáhá malým a středním podnikům k přístupu na globální trhy pomocí sdružování za účelem výměny myšlenek a dobré praxe. Sdílení znalostí pomáhá inovacím a poskytuje všem partnerům schopnost přizpůsobit se změnám trhu. Spolupráce je klíčem k rozvoji DBE a ke změnám v myšlení firem směrem k vyšší úrovni spolupráce. Hlavní rysy DBE jsou znázorněny na obr. 15.



Obr. 15 Čtyři hlavní rysy DBE, ERISA (2007)

Při postupném zavádění se DBE stává samoorganizujícím a samoudržitelným systémem. Ze začátku ovšem jeho implementace vyžaduje systémový přístup zahrnující systém regionálního průmyslu, výzkumu a vývoje a místní správy. Tento proces vyplývá z interakcí mezi různými vlastníky a producenty znalostí (jako jsou velké a malé výrobní podniky, univerzity, výzkumná a vývojová centra, regionální činitelé, profesní sdružení, podnikatelské asociace a organizace pro podporu podnikání). Role veřejné správy, EVROPSKÁ KOMISE (2007), může mít formu podpory v oblastech jako je dodržování zákonnosti a vyváženosti zájmů, kultura komunikace, důvěryhodnost a záruky, organizace a synchronizace, patenty a regulační rámce, standardy technologií a infrastruktury. DBE má význam i pro orgány územní samosprávy, které se zabývají regionálními operačními programy. Postavení veřejné správy v rozvoji DBE vyplývá i z obr. 16.



Obr. 16 DBE jako model pro místní rozvoj, ERISA (2007)

Díky 6. rámcovému programu EU byla tato technologie v rámci pilotních projektů implementována a dále šířena v podnikatelských strukturách tří evropských regionů (španělská Aragonie, anglické West Midlands a finské Tampere), neboť DBE je model také pro regionální rozvoj. V regionálním kontextu může implementace DBE napomoci evropským MSP soutěžit na světové úrovni bez ohledu na to, jak malé, vzdálené nebo neznámé jsou. To přímo ovlivňuje ekonomický vývoj MSP a tudíž i sociální a ekonomický rozvoj regionu jako celku. Regionální management má roli katalyzátoru urychlujícího vývoj DBE. Tato role se projevuje

např. tím, ERISA (2007), EVROPSKÁ KOMISE (2007), že pomůže spolufinancovat vytvoření základní regionální infrastruktury pro DBE tím, že její vytvoření nasměrují na místní IT firmy apod.). DBE může také propojit e-Government s MSP nebo propagovat znalosti „zabudované“ v místním teritoriu.

Regionální management však může být, kromě své primární zastřešující role, i prvkem - uživatelem DBE, který využívá i poskytuje svoje znalosti. Modely hodnocení spokojenosti občanů navrhované v této práci mohou při své aplikaci a zařazení do struktury DBE poskytovat znalosti, které může využívat více subjektů v regionu, včetně MSP. Některé orgány územní samosprávy se již na volném poskytování znalostí podílejí. Např. Městský úřad v Banské Bystrici postupně otestoval a nasadil do běžného provozu několik řešení na bázi Open Source, a to jak na straně serverové infrastruktury, tak na straně klientů. Využívání otevřené standardy jako např. LINUX, IMAP, OpenDocument. Město vyvinulo vlastní GIS portál (na bázi Open Source systémů) pro zveřejnění územního plánu s jeho změnami i doplňky. Má dobrou technickou podporu, jeho užívání je zdarma včetně aktualizací, kód je otevřený s možností rozšíření, systém je nezávislý na webovém prohlížeči a nevyžaduje na straně klienta žádný komerční software. Tento přístup byl prezentován na letošní Konferenci ISSS v Hradci Králové i s nabídkou na výměnu zkušeností a know-how, protože Open Source technologie již testuje a nasazuje mnoho samospráv. TUHÁRSKY - VOSKÁR (2008)

5 Návrh modelů klasifikačního systému

Vstupem modelů systému klasifikace občanů do tříd podle míry jejich spokojenosti jsou data ve formě tabulky v aplikaci MS Excel, exportované z výchozí tabulky MS Access. Jsou to výsledky dotazníkového šetření Společné evropské indikátory - indikátor A1 (Spokojenost občanů s místní společenstvím - Obecná spokojenost občanů s různými rysy samosprávy), poskytnuté Městským úřadem Chrudim. Metodika šetření je uvedena v Příloze 2: Metodika dotazníkového průzkumu, TIMUR (2006).

Během samotného modelování je vytvářeno více modelů. V používaných nástrojích je množství parametrů, které jsou měněny. Proto je nutné důkladně zaznamenávat nastavené hodnoty včetně zdůvodnění, proč je vybrána zrovna tato kombinace nastavených parametrů. CHAPMAN (2000). Zaznamenávání všech parametrů usnadňuje orientaci v modelech a datech a může zabránit opakování operací.

5.1 Formalizace klasifikační úlohy a návrh atributů

Klasifikační úlohu o spokojenosti občanů můžeme formalizovat na základě obecné formalizace klasifikační úlohy uvedené v části 3.3. Zdrojem dat úlohy jsou výsledky dotazníkového průzkumu. Dotazováno bylo 701 občanů, kteří odpovídali na 94 dotazníkových otázek. Jednotlivé respondenty můžeme z hlediska klasifikační úlohy považovat za objekty klasifikace a dotazníkové otázky za jejich atributy. Z nich bylo pro další zpracování vybráno 73 necílových atributů. 21 otázek nebylo z různých důvodů zahrnuto - některé představují pouze pořadí atributů, jiné měly příliš velké množství chybějících hodnot. Přehled atributů obsahující cílový atribut (s24) a 73 necílových atributů vybraných pro další analýzu je uveden v tabulce 2. Název atributu odpovídá původnímu značení zdrojových dat a odpovídá dotazníku, který je uveden v Příloze 1. V něm lze vyhledat přesné znění otázky. Vybraná data jsou uložena v matici $\boldsymbol{M}$, tvořené n řádky a m sloupci, kde $n = 701$ a $m = 73$,

$$\boldsymbol{M} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{11} & \cdots & x_{nm} \end{bmatrix}. \tag{5.1}$$

Řádky matice reprezentují sledované objekty - respondenty, *i*-tý respondent je tedy řádek $\boldsymbol{x}_i$, kde $m = 73$,

$$\boldsymbol{x}_i = [x_{i1}, x_{i2}, \ldots, x_{im}]. \tag{5.2}$$

Sloupce datové matice odpovídají atributům - dotazníkovým otázkám, *j*-tý atribut (*j*-tý sloupec) je značen symbolem $\boldsymbol{S}_j$, kde $n = 701$,

$$\boldsymbol{S}_j = [x_{1j}, x_{2j}, \ldots, x_{nj}]^T . \tag{5.3}$$

Cílový atribut $\boldsymbol{Z}$, kde $n = 701$.

$$\boldsymbol{Z}: [y_1, y_2, \ldots, y_n]^T . \tag{5.4}$$

je atribut s24 „spokojenost s kvalitou okolního životního prostředí“ (viz tab. 2).

Trénovací data jsou uložena v matici $\boldsymbol{M}_{TR}$, tvořené *n* řádky a *m* sloupci, kde *n* odpovídá náhodnému výběru trénovacích příkladů a *m* odpovídá počtu vybraných atributů,

$$\boldsymbol{M}_{TR} = \begin{bmatrix} x_{11} & x_{12} & \ldots & x_{1m} & y_1 \\ x_{21} & x_{22} & \ldots & x_{2m} & y_2 \\ \vdots & \vdots & & \vdots & \vdots \\ x_{n1} & x_{11} & \ldots & x_{nm} & y_n \end{bmatrix}. \tag{5.5}$$

Objekt - respondent (trénovací příklad) z této matice je značen

$$\boldsymbol{r}_i = [\boldsymbol{x}_i, y_i]. \tag{5.6}$$

Trénovací data můžeme tedy považovat za množinu respondentů

$$\boldsymbol{M}_{TR} = \{\boldsymbol{r}_i, i = 1, \ldots, n\}, \tag{5.7}$$

kde *n* odpovídá počtu trénovacích příkladů získaných náhodným výběrem.

Tab. 2 Přehled vybraných atributů

Označení	Atribut	Označení	Atribut	Označení	Atribut
s1	pohlaví	s26	8	s51	D - social_služby
s2	věk	s27	9	s52	D - social_bydleni
s3	zaměstnání	s28	10	s53	D - policie
s4	adresa	s29	11	s54	D - školy
s5	kuřák	s30	12	s55	D - veřejná doprava
s6	1a	s31	13a	s56	E - parky
s7	2	s32	13b	s57	E - zastavěné
s8	3	s33	13c	s58	E - odpady
s9	4	s34	13d	s59	E -ovzduší
s10	4zdrav	s35	13e	s60	E - hluk noc
s11	4social	s36	13f	s61	E - hluk den
s12	4ms,zs	s37	13g	s62	F - rekvalifikace
s13	4zs	s38	13h	s63	F - podnikáni
s14	4MHD	s39	13ch	s64	F - nezaměstnanost
s15	4TS	s40	B - doma	s65	F - hospodaření
s16	4MP	s41	B - noc	s66	F - kouření restaurace
s17	5	s42	B - ulice	s67	F - kouření nádraží
s18	5a	s43	B - prostranství	s68	F - kouření úřady
s19	5b	s44	C - sport	s69	F - kouření školy
s20	5c	s45	C - divadlo	s70	F - kouření ostatní
s21	5d	s46	C - muzea	s71	G - plánování
s22	5e	s47	C - kultur	s72	G - NNO
s23	5f	s48	C - knihovna	s73	G - žádosti
s24	**6**	s49	D - lékaři	s74	G - volby
s25	7	s50	D - nemocnice		

Zařazování atributů se silnou korelací zbytečně komplikuje modely, proto se silně korelované atributy společně do modelů nezařazují. Nejvyšší hodnoty pro Pearsonův koeficient korelace dvojic atributů jsou uvedeny v následující tab. 3.

Tab. 3 Pearsonův koeficient korelace

Označení	Atribut	Označení	Atribut	Korelace
s2	věk	s3	zaměstnání	0,806
s68	F - kouření úřady	s69	F - kouření školy	0,771
s42	B - ulice	s43	B - prostranství	0,755
s71	G - plánování	s72	G - NNO	0,743
s66	F - kouření restaurace	s67	F - kouření nádraží	0,719
s62	F - rekvalifikace	s63	F - podnikáni	0,711
s67	F - kouření nádraží	s70	F - kouření ostatní	0,693
s73	G - žádosti	s74	G - volby	0,693
s66	F - kouření restaurace	s70	F - kouření ostatní	0,686
s67	F - kouření nádraží	s68	F - kouření úřady	0,671

5.2 Popis dat

Podrobný popis dat vybraných pro další analýzu - datový slovník - je uveden v tab. 4. Atributy s1 - s5 zahrnují data socio-demografická. Ve zbývajících řádcích jsou potom odpovědi na otázky o spokojenosti, tzn. hodnocení, která respondenti přiřadili jednotlivým aspektům života ve městě v závislosti na své spokojenosti s nimi. Dotazník má v některých částech hierarchickou strukturu - otázky jsou dále členěny podotázkami (s9, s17). Stupnice pro hodnocení je stanovena v relativních jednotkách:

- atribut s6 spokojenost s místním společenstvím, sestupně 4 - 1 (1 velmi spokojen, 2 mírně spokojen, 3 mírně nespokojen, 4 velmi nespokojen);
- další atributy, není-li uvedeno jinak, vzestupně 0 - 10 (0 velmi nespokojen - 10 velmi spokojen).

Tab. 4 Datový slovník

Označení	Význam atributu	Hodnoty
s1	pohlaví (muž, žena)	1 / 2
s2	věk	[16, 82]
s3	zaměstnání (student, zaměstnaný, nezaměstnaný, důchodce)	[1, 4]
s4	adresa (městské části Chrudimě)	[1, 9]
s5	kuřák (počet cigaret denně)	[0, 40]
s6	spokojenost s místním společenstvím	[1, 4]
s7	mezilidské vztahy	[0,10]
s8	možnost provozovat záliby a zájmy	[0,10]
s9	veřejné služby	[0,10]
s10 - s16	spokojenost se základními veřejnými službami	[0,10]
s17	hodnocení spokojenosti s fungováním MÚ	[0,10]
s18 - s23	hodnocení spokojenosti s fungováním MÚ	[0,10]
s24	**kvalita okolního životního prostředí**	[0,10]
s25	možnost zaměstnání ve městě	[0,10]
s26	zapojení veřejnosti - strategické plánování	[0,10]
s27	Jste dlouhodobě léčen nebo sledován lékařem? (ano, ne , bez odpovědi)	[1, 3]
s28	počet návštěv zdravot. zařízení v Chrudimi za posledních 12 měsíců: (0; 1 až 5; 6 až 10; 11 a více)	[1, 4]
s29	počet hodin sportu týdně (0; 1 až 3; 4 až 7; 8 až 10; 11 až 15; 16 až 20; více než 20)	[1, 7]
s30	Domníváte se, že prostředí a životní podmínky ovlivňují Vaše zdraví? (velmi, spíše ano, nevím, spíše ne, určitě ne)	[1, 4]
s31 - s39	hodnocení spokojenosti - každodenní život v obci	[0,10]
s40 - s43	hodnocení bezpečnosti	[0,10]
s44 - s48	hodnocení kvality veřejných služeb	[0,10]
s49 - s55	hodnocení dostupnosti základních služeb	[0,10]
s56 - s61	hodnocení kvality složek městského prostředí	[0,10]
s62 - s70	hodnocení možností ve městě	[0,10]
s71 - s74	možnosti zapojení do rozhodování ve městě	[0,10]

5.3 Příprava dat

Příprava dat v MS Excel představovala zjištění popisných charakteristik dat (hodnoty, minimum, maximum, četnosti, datový typ, korelace atributů apod.), ošetření chybějících hodnot, ošetření zjevně chybných hodnot a převod hodnot atributu s4 - adresa na název městské

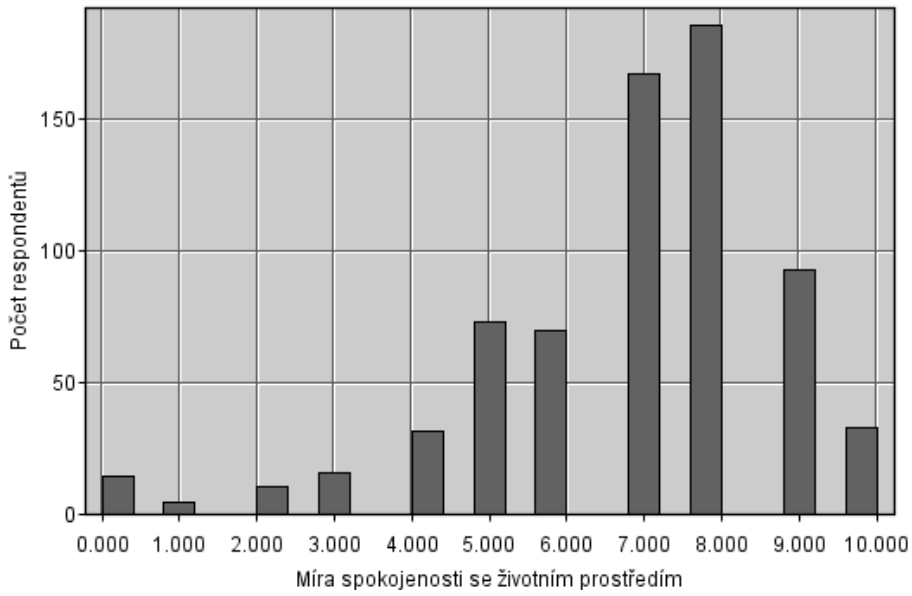
části (zaznamenáván tazateli nejednotně, např. jméno ulice, místní názvy apod.). Datový soubor byl z listu MS Excel exportován do formátu *.csv a načten do Clementine. Clementine zahrnuje jednoduché a účinné nástroje pro předzpracování dat, proto zde byly provedeny další úpravy vstupního souboru - jak pro modelování v Clementine, tak i pro MATLAB.

FIS i jednotlivé algoritmy rozhodovacích stromů mají odlišné požadavky na data, proto byla příprava dat pro modelování prováděna samostatně pro každou skupinu modelů. Podle potřeby byly prováděny následující typy úprav:

- kategorizace atributů s velkým množstvím hodnot do menšího počtu intervalů (lepší orientace v rozhodovacím stromu, tvorba funkcí příslušnosti a pravidel pro FIS, kategorizace cílového atributu do 2 hodnot spokojen - nespokojen);
- reklasifikace číselného značení na slovní (lepší orientace v rozhodovacím stromu);
- vyloučení objektů se zjevně chybnými hodnotami nebo s hodnotami chybějícími (pro FIS);
- náhodné rozdělení vstupní datové matice na trénovací a testovací objekty.

5.4 Modelování spokojenosti - rozhodovací stromy

Cílem zvolené klasifikační úlohy je zařazení občanů do tříd podle míry jejich spokojenosti. Jak je uvedeno v části 2.1, úkolem je nalezení vhodných socio-demografických atributů a aspektů spokojenosti občanů a na jejich základě určit míru jiného (cílového) aspektu spokojenosti. Jako cílový atribut byl vybrán atribut s24 - spokojenost občanů s kvalitou okolního životního prostředí. Hodnoty tohoto atributu byly rozděleny do dvou intervalů hodnot (spokojen - nespokojen) a v některých modelech byl vyrovnán počet objektů s hodnotou „nespokojen“, neboť ta se vyskytuje méně často než „spokojen“ (závisí na stanovení mezí intervalů). Četnosti výskytu jednotlivých hodnot cílového atributu jsou na obr. 17.



Obr. 17 Histogram cílového atributu

V modelech je pro větší přehlednost dále využíváno zkrácené značení atributů uvedené v tab. 2 - Přehled vybraných atributů. To odpovídá původnímu značení v dotazníku a lze podle něj v Příloze 1: Dotazník vyhledat přesné znění otázek. Během procesu tvorby a optimalizace modelů bylo odzkoušeno velké množství modifikací. Mezi modifikace, které nejvíce ovlivňují správnost klasifikace, patří kombinace vstupních atributů, výběr algoritmů a nastavení jejich parametrů. Pro algoritmus C5.0 lze nastavit metody boosting a křížová validace, které používají vícenásobnou aplikaci modelu na chybně zařazené objekty. Vybranými algoritmy pro vytváření rozhodovacích stromů při modelování spokojenosti s kvalitou životního prostředí jsou algoritmy C5.0, CHAID, C&RT a QUEST v prostředí Clementine. V následující části práce je proto uveden přehled pouze těch modelů, které dosahují nejvyšší správnosti klasifikace. Přehled vybraných modelů a jejich vstupních atributů je uveden v tab. 5.

Vstupní datová matice byla náhodně rozdělena na trénovací a testovací objekty v poměru 75 % trénovacích a 25 % testovacích. Hodnoty atributů byly sloučeny do menšího počtu intervalů. Numerické atributy byly označeny slovně.

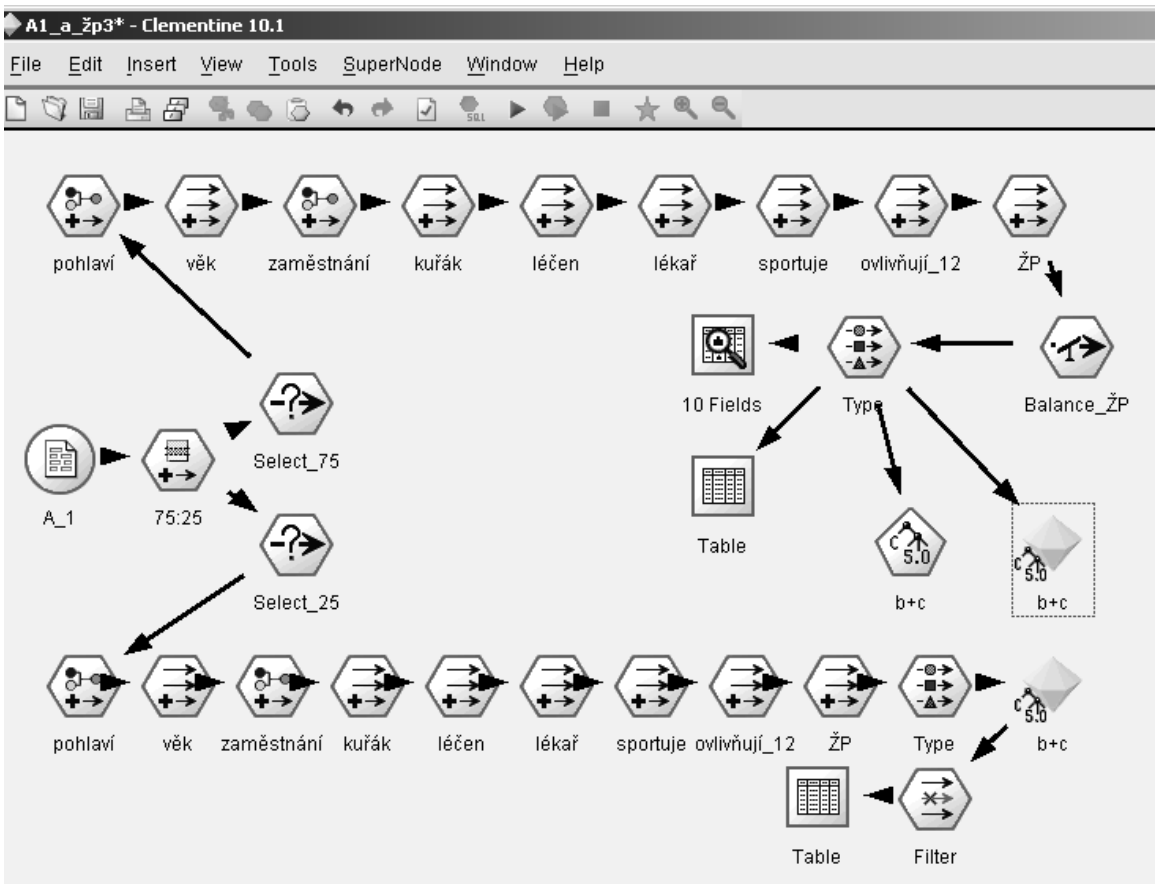
Model DT1

Model s algoritmem C5.0 s využitím metod boosting a křížová validace s nastavením na 20 cyklů. V modelu byl vyrovnán počet objektů s hodnotou cílového atributu „nespokojen“ [0, 5]. Chybějící hodnoty nebyly nahrazeny, algoritmus s nimi pracuje jako s další hodnotou.

Vstupní atributy tohoto modelu jsou „objektivní“ socio-demografické atributy, které se netýkají subjektivního názoru na kvalitu života ve městě:

- pohlaví (s1), věk (s2), zaměstnání (s3), adresa (s4);
- počet denně vykouřených cigaret (s5);
- „Jste dlouhodobě léčen?“ (s27);
- počet návštěv zdravot. zařízení v Chrudimi za posledních 12 měsíců (s28);
- počet hodin sportu týdně (s29);
- „Domníváte se, že prostředí a životní podmínky ovlivňují Vaše zdraví?“ (s30).

Celý postup v Clementine (stream) je znázorněn na obr. 18. Model byl vytvořen z trénovacích dat (horní část streamu). Výsledný model (na obrázku označen „b+c“ a zarámován) je pak testován na testovacích datech (dolní část streamu). Rozhodovací strom (obr. 19) je zobrazen tak, že za šipkou je uvedeno zařazení do třídy, v závorce je vždy uveden počet příkladů pokrytých danou větví a správnost klasifikace. Hodnocení správnosti všech modelů je v části 5.6 - Vyhodnocení výsledků.



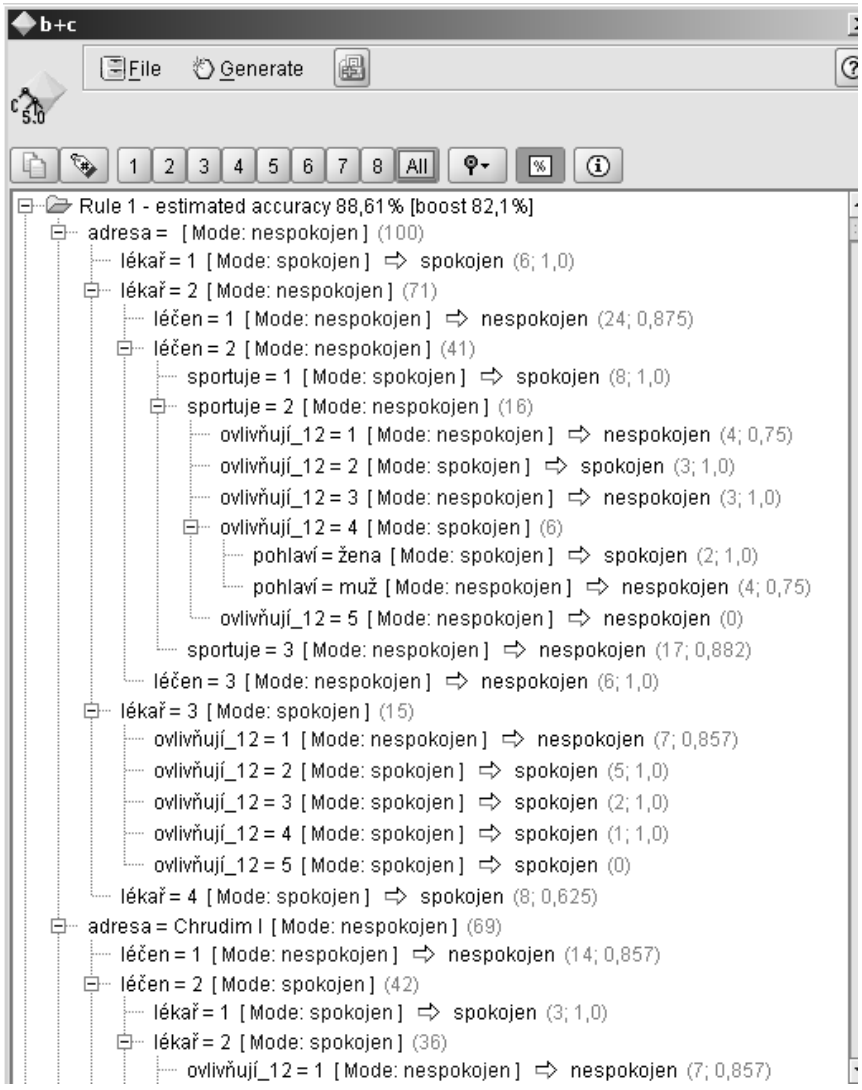
Obr. 18 Stream pro tvorbu modelu DT1

Model DT2

Modifikace modelu DT1. V modelu byl přibližně vyrovnán počet objektů pro hodnoty všech vstupních atributů. Ostatní parametry modelu byly zachovány.

Model DT3

Modifikace modelu DT1. V modelu byl přibližně vyrovnán počet objektů pro hodnoty všech vstupních atributů. Dále byl rozšířen interval hodnoty cílového atributu „nespokojen“ o jednu hodnotu z [0, 5] na [0, 6]. Ostatní parametry modelu byly zachovány.



Obr. 19 Náhled části rozhodovacího stromu modelu DT1

Model DT4

Model s algoritmem C5.0 s nastavením metod boosting a křížová validace. V modelu byl vyrovnán počet objektů s hodnotou cílového atributu „nespokojen“ [0, 5]. Jako vstupní atributy tohoto modelu jsou použity všechny atributy.

Model DT5

Model s algoritmem CHAID s nastavením hloubky stromu do 20. úrovně. V modelu byl vyrovnán počet objektů s hodnotou cílového atributu „nespokojen“ [0, 5]. Jako vstupní atributy tohoto modelu jsou použity všechny atributy.

Model DT6

Model s algoritmem QUEST s nastavením hloubky stromu do 20. úrovně. V modelu byl vyrovnán počet objektů s hodnotou cílového atributu „nespokojen“ [0, 5]. Jako vstupní atributy tohoto modelu jsou použity všechny atributy.

Model DT7

Model s algoritmem C&RT s nastavením hloubky stromu do 20. úrovně. V modelu byl vyrovnán počet objektů s hodnotou cílového atributu „nespokojen“ [0, 5]. Jako vstupní atributy tohoto modelu jsou použity všechny atributy.

5.5 Modelování spokojenosti - fuzzy inferenční systémy

Úkol modelování spokojenosti pomocí fuzzy inferenčního systému typu Mamdani zůstává stejný jako u klasifikace metodami rozhodovacích stromů - nalézt vhodné socio-demografické atributy a aspekty spokojenosti občanů a na jejich základě určit míru spokojenosti s kvalitou okolního životního prostředí. Výběr vstupních atributů byl proveden dvěma způsoby - výběr pomocí rozhodovacích stromů v Clementine ze všech atributů a dále pak vlastní výběr „objektivních“ socio-demografických atributů, které charakterizují respondenta. V druhém případě byly záměrně vybrány popisné atributy, které se netýkají subjektivního názoru na kvalitu života ve městě (např. věk, frekvence návštěv lékaře, kouření apod.). Závažným omezením fuzzy modelů je skutečnost, že pro návrh FIS lze použít pouze ordinální vstupní atributy. Nelze tedy zařadit atributy, které by byly pro potřeby městského či regionálního managementu jistě významné (např. pohlaví, adresa, zaměstnání apod.). I v případě fuzzy modelů bude pro větší přehlednost dále využíváno zkrácené značení uvedené v tab. 2 - Přehled vybraných atributů.

Fuzzy Logic Toolbox simulačního prostředí MATLAB umožňuje naplnění znalostní a datové báze fuzzy systému i následné vyhodnocování vytvořených modelů. Práce je možná s využitím grafického uživatelského interface, ale při větším množství pravidel systém

pracuje pomalu. K modelování a vyhodnocování modelů bylo proto využito prostředí příkazového okna.

Při návrhu fuzzy modelu je třeba definovat:

- vstupy, jejich jména a rozsah;
- vstupní funkce příslušnosti, jejich jména, tvar a polohu v univerzu;
- výstupy, jejich jména a rozsah;
- výstupní funkce příslušnosti, jejich jména, tvar a polohu v univerzu;
- fuzzy inferenční pravidla.

Všechny informace o navrhovaném FIS obsahuje FIS matice, což je objekt prostředí MATLAB, uložený ve speciálním textovém formátu (*.fis). Ze tří možných formátů zápisu pravidel - slovního, komprimovaného a symbolického - byl zvolen komprimovaný formát. Tato forma je zkráceným zápisem pravidla a užívá místo skutečných jmen proměnných a jejich slovních hodnot pouze jejich pořadová čísla. Např. pravidlo

IF vstup1 je mf1 AND vstup2 je mf1 THEN výstup je mf3 s vahou 1 (5.8)

kde mf je funkce příslušnosti, zapíšeme v komprimovaném formátu ve tvaru

1 1, 3 (1) : 1 (5.9)

kde číslo za čárkou znamená konsekvent, číslo v závorce je váha pravidla a 1 za dvojtečkou znamená spojku AND (2 je OR).

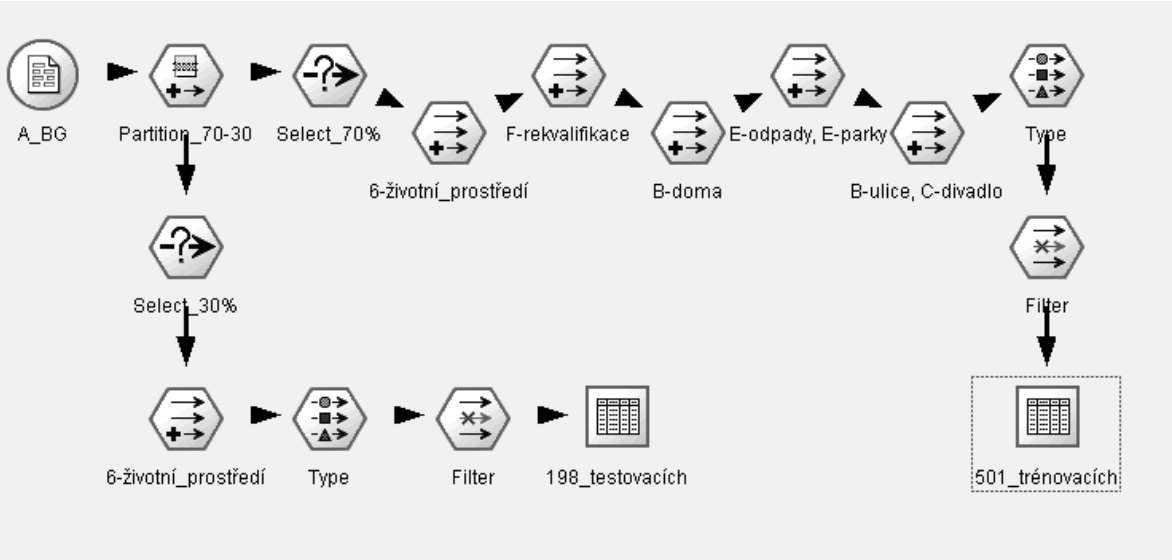
Naplnění znalostní báze je klíčovým momentem tvorby FIS a lze provést třemi základními způsoby - převodem znalostí experta, intuitivně (s pomocí popisu systému) a automatickým generováním. Tyto způsoby je možné kombinovat. JURA (2003), NOVÁK (2000). V této práci byl použit postup tvorby znalostní báze přímo z trénovacích dat - generováním pravidel a funkcí příslušnosti z intervalů vytvořených slučováním hodnot atributů. Jedna část modelů pak byla vytvořena zařazením všech generovaných pravidel, pro druhou část modelů byl využit intuitivní přístup k redukci generovaných pravidel s přihlédnutím k četnostem hodnot jednotlivých atributů. Z „objektivních“ socio-demografických atributů byly vybrány:

- věk (s2);
- počet denně vykouřených cigaret (s5);
- počet návštěv zdravot. zařízení v Chrudimi za posledních 12 měsíců (s28);
- počet hodin sportu týdně (s29);
- „Domníváte se, že prostředí a životní podmínky ovlivňují Vaše zdraví?“ (s30).

Při výběru vstupních atributů v Clementine byly nejprve pomocí algoritmů C5.0 a CHAID vytvořeny rozhodovací stromy ze všech atributů. To umožní získat představu o důležitosti jednotlivých atributů - nejdůležitější atributy jsou nejblíže kořene stromu (mají největší informační přínos). Výsledkem analýzy byl výběr šesti atributů, které mají největší vliv na cílový atribut - spokojenost se životním prostředím. Jsou to následující atributy, uvedené v pořadí od nejdůležitějších (přesné znění otázky odpovídající jednotlivým atributům je v Příloze 1: Dotazník):

1. E - odpady (s58);
2. B - ulice (s42), E - parky (s56);
3. F - rekvalifikace (s62), B - doma (s40), C - divadlo (s45).

Redukovaná vstupní datová matice šesti vybraných atributů byla v Clementine náhodně rozdělena na trénovací a testovací objekty v poměru 70 % trénovacích a 30 % testovacích. Hodnoty atributů trénovacích objektů byly (podle počtu hodnot, které nabývají) sloučeny do dvou nebo třech intervalů, které se staly základem pro tvorbu funkcí příslušnosti a tvorbu pravidel. Trénovací data byla exportována do MS Excel a upravena do komprimovaného formátu pravidel pro FIS typu Mamdani. Hodnoty testovacích dat byly v nezměněné podobě využity pro vyhodnocení modelů (viz část 5.6). Na obr. 20 je znázorněn stream v Clementine pro tuto část.



Obr. 20 Rozdělení dat a úprava vybraných atributů v Clementine

Během procesu tvorby a optimalizace modelů bylo odzkoušeno velké množství modifikací všech částí FIS. Mezi modifikace, které nejvíce ovlivňují správnost klasifikace patří kombinace vstupních atributů, nastavení parametrů a určení počtu vstupních i výstupních

funkcí příslušnosti a sestavení báze pravidel. V následující části práce je proto uveden přehled pouze těch modelů, které dosahují nejvyšší správnosti klasifikace. Přehled vybraných modelů a jejich vstupních atributů je uveden v tab. 5. Je-li popisován částečně modifikovaný model, je z důvodu omezeného prostoru uveden skript pouze modifikované části a není uváděna ani báze pravidel. Báze pravidel je pro ilustraci uvedena jen u prvního modelu FIS1.

Tab. 5 Přehled modelů a jejich vstupních atributů

Označení atributu	FIS1	FIS2	FIS3	FIS4	FIS5
E - odpady (s58)	X	X	X	X	
B - ulice (s42)	X	X	X	X	
E - parky (s56)	X	X	X	X	
F - rekvalifikace (s62)				X	
B - doma (s40)				X	
C - divadlo (s45)				X	
věk (s2)					X
kuřák (s5)					X
u lékaře (s28)					X
sportuje (s29)					X
Ovlivňuje prostředí zdraví? (s30)					X

Model FIS1

Výběr vstupních atributů tohoto modelu byl proveden pomocí rozhodovacích stromů, vybrány byly tři nejdůležitější atributy. Znalostní báze byla vytvořena intuitivně - redukcí sady všech generovaných pravidel s přihlédnutím k četnostem hodnot jednotlivých atributů. Z tohoto důvodu je omezen počet vstupů na tři. Větší počet kombinací, který vzniká zvýšením počtu vstupních funkcí příslušnosti, nelze bez experta smysluplně interpretovat a transformovat do znalostní báze. Počet hodnot cílového atributu je redukován rozdělením do dvou intervalů (spokojen - nespokojen). Úplná definice FIS1 je v následujícím skriptu.

```
[System]
Name='A1_int'
Type='mamdani'
NumInputs=3
NumOutputs=1
NumRules=9
AndMethod='min'
OrMethod='max'
ImpMethod='min'
AggMethod='max'
DefuzzMethod='centroid'
```

```
[Input1]
Name='Eodpady'
Range=[0 10]
NumMFs=3
MF1='nespok':'trimf',[0 0 7]
MF2='stred':'trimf',[0 7 10]
MF3='spok':'trimf',[7 10 10]

[Input2]
Name='Eparky'
Range=[0 10]
NumMFs=3
MF1='nespok':'trimf',[0 0 7]
MF2='stred':'trimf',[0 7 10]
MF3='spok':'trimf',[7 10 10]

[Input3]
Name='Bulice'
Range=[0 10]
NumMFs=3
MF1='nespok':'trimf',[0 0 6.5]
MF2='stred':'trimf',[0 6.5 10]
MF3='spok':'trimf',[6.5 10 10]

[Output1]
Name='ZP'
Range=[0 3]
NumMFs=2
MF1='nespokojen':'trimf',[0 1 1.51]
MF2='spokojen':'trimf',[1.51 2 3]

[Rules]
3     3     2     ,     2     (1)   :     1
1     1     1     ,     1     (1)   :     1
1     3     2     ,     2     (1)   :     1
2     3     2     ,     2     (1)   :     1
2     1     1     ,     1     (1)   :     1
1     1     2     ,     1     (1)   :     1
3     1     1     ,     1     (1)   :     1
3     3     1     ,     2     (1)   :     1
3     3     3     ,     2     (1)   :     1
```

Model FIS2

Tento model vychází z předchozího, pro dosažení větší správnosti klasifikace má rozšířenu bázi pravidel.

```
[System]
Name='A1_int_rul18'
Type='mamdani'
NumInputs=3
NumOutputs=1
NumRules=18
AndMethod='min'
OrMethod='max'
ImpMethod='min'
AggMethod='max'
DefuzzMethod='centroid'
```

Model FIS3

Výběr vstupních atributů tohoto modelu byl proveden pomocí rozhodovacích stromů, vybrány byly tři nejdůležitější atributy. Znalostní báze byla vytvořena intuitivně jako u předchozích modelů. Hodnoty cílového atributu jsou rozděleny do dvou intervalů (spokojen - nespokojen). Model má jiné tvary vstupních i výstupních funkcí příslušnosti, použita je i jiná defuzzifikační metoda. Úplná definice FIS3 je uvedena v následujícím skriptu.

```
[System]
Name='A1_int_singleton_gauss'
Type='mamdani'
Version=2.0
NumInputs=3
NumOutputs=1
NumRules=9
AndMethod='min'
OrMethod='max'
ImpMethod='min'
AggMethod='max'
DefuzzMethod='lom'

[Input1]
Name='Eodpady'
Range=[0 10]
NumMFs=3
MF1='nespok':'gaussmf',[3 0]
MF2='stred':'gaussmf',[2 5]
MF3='spok':'gaussmf',[3 10]

[Input2]
Name='Eparky'
Range=[0 10]
NumMFs=3
MF1='nespok':'gaussmf',[3 0]
MF2='stred':'gaussmf',[2 5]
MF3='spok':'gaussmf',[3 10]

[Input3]
Name='Bulice'
Range=[0 10]
NumMFs=3
MF1='nespok':'gaussmf',[3 0]
MF2='stred':'gaussmf',[2 5]
MF3='spok':'gaussmf',[3 10]

[Output1]
Name='ZP'
Range=[1 2]
NumMFs=2
MF1='nespokojen':'trimf',[1 1 1]
MF2='spokojen':'trimf',[2 2 2]

[Rules]
...
```

Model FIS4

Výběr vstupních atributů tohoto modelu byl proveden pomocí rozhodovacích stromů, vybráno bylo tentokrát šest nejdůležitějších atributů. Znalostní báze byla vytvořena ze všech 501 generovaných pravidel. Počet hodnot cílového atributu je redukován rozdělením do dvou intervalů (spokojen - nespokojen). Úplná definice FIS4 je uvedena v následujícím skriptu.

```
[System]
Name='A1'
Type='mamdani'
NumInputs=6
NumOutputs=1
NumRules=501
AndMethod='min'
OrMethod='max'
ImpMethod='min'
AggMethod='max'
DefuzzMethod='centroid'

[Input1]
Name='Bdoma'
Range=[0 10]
NumMFs=3
MF1='nespok':'trimf',[0 0 7]
MF2='stred':'trimf',[0 7 10]
MF3='spok':'trimf',[7 10 10]

[Input2]
Name='Bulice'
Range=[0 10]
NumMFs=3
MF1='nespok':'trimf',[0 0 6.5]
MF2='stred':'trimf',[0 6.5 10]
MF3='spok':'trimf',[6.5 10 10]

[Input3]
Name='Cdivadlo'
Range=[0 10]
NumMFs=3
MF1='nespok':'trimf',[0 0 6.5]
MF2='stred':'trimf',[0 6.5 10]
MF3='spok':'trimf',[6.5 10 10]

[Input4]
Name='Eodpady'
Range=[0 10]
NumMFs=3
MF1='nespok':'trimf',[0 0 7]
MF2='stred':'trimf',[0 7 10]
MF3='spok':'trimf',[7 10 10]

[Input5]
Name='Eparky'
Range=[0 10]
NumMFs=3
MF1='nespok':'trimf',[0 0 7]
MF2='stred':'trimf',[0 7 10]
MF3='spok':'trimf',[7 10 10]
```

```
[Input6]
Name='Frekvalifikace'
Range=[0 10]
NumMFs=3
MF1='nespok':'trimf',[0 0 6.5]
MF2='stred':'trimf',[0 6.5 10]
MF3='spok':'trimf',[6.5 10 10]

[Output1]
Name='ZP'
Range=[0 3]
NumMFs=2
MF1='nespokojen':'trimf',[0 1 1.51]
MF2='spokojen':'trimf',[1.51 2 3]

[Rules]
...
```

Model FIS5

Výběr vstupních atributů tohoto modelu byl proveden z ordinálních „objektivních“ socio-demografických atributů, které se netýkají subjektivního názoru na kvalitu života ve městě. Znalostní báze byla vytvořena intuitivně - výběrem ze všech možných kombinací vstupních i výstupních funkcí příslušnosti s přihlédnutím k četnostem hodnot jednotlivých atributů. Hodnoty cílového atributu byly ponechány v původním rozsahu. Úplná definice FIS5 je uvedena v následujícím skriptu.

```
[System]
Name='A1_demo_40rul_zp'
Type='mamdani'
NumInputs=5
NumOutputs=1
NumRules=40
AndMethod='min'
OrMethod='max'
ImpMethod='min'
AggMethod='max'
DefuzzMethod='centroid'

[Input1]
Name='vek'
Range=[15 120]
NumMFs=3
MF1='mlady':'trimf',[15 15 40]
MF2='stredni':'trimf', [25 40 70]
MF3='stary':'trapmf', [40 70 120 120]

[Input2]
Name='kurak'
Range=[0 100]
NumMFs=2
MF1='nekurak':'trimf',[0 0 0.5]
MF2='kurak':'trapmf', [0.5 20 100 100]
```

```
[Input3]
Name='u_lekare'
Range=[1 4]
NumMFs=2
MF1='malo':'gaussmf',[1.2 1]
MF2='casto':'gaussmf',[1.2 4]

[Input4]
Name='sportuje'
Range=[1 6]
NumMFs=2
MF1='malo':'gaussmf',[1.8 1]
MF2='casto':'gaussmf',[1.8 6]

[Input5]
Name='ovlivnuje'
Range=[1 5]
NumMFs=3
MF1='velmi':'gaussmf',[1.5 1]
MF2='nevim':'gaussmf',[0.3 3]
MF2='ne':'gaussmf',[1.5 5]

[Output1]
Name='spokojensZP'
Range=[0 10]
NumMFs=2
MF1='nespokojen':'gaussmf',  [0.4 -1.11e-016]
MF2='spokojen':'gaussmf',  [0.4 10]

[Rules]
...
```

5.6 Vyhodnocení výsledků

Pro hledání znalostí pro potřeby klasifikace bylo postupováno pomocí metod učení s učitelem. Metody hodnocení modelů jsou tedy založeny na testování shody nalezených znalostí s informací učitele. Pro testování existuje řada metod podle toho, jaká data použijeme pro učení a jaká pro testování - testování v celých trénovacích datech, křížová validace, leave-one-out, bootstrap a testování na testovacích datech. Modely jsou v této práci testovány na testovacích datech.

Důvodem testování správnosti modelu na datech, která klasifikátor „nezná“ je ten, že pokud je množina dat příliš velká, dosáhne algoritmus malé chyby za cenu, že vytvořený rozhodovací strom bude příliš „košatý“. Tento jev se nazývá přeučení (overfitting) a vede k malé až žádné schopnosti modelu předpovídat na nových datech. ŽELEZNÝ - KLÉMA - ŠTĚPÁNKOVÁ (2003).

Při testování zjišťujeme, v kolika případech se zařazení objektů klasifikátorem shoduje se zařazením učitele a kolikrát se dopustil chyby. Údaje o správnosti klasifikačního modelu zachycujeme do matice záměn (confusion matrix). Tato matice dává do souvislosti výsledky

klasifikace dat s informací o správném zařazení těchto příkladů do tříd. V tab. 6 je porovnání klasifikace do dvou tříd, „+“ a „-“, tzv. čtyřpolní tabulka. Ve sloupcích je uvedena klasifikace systémem, v řádcích pak správná klasifikace. *TP* (správně pozitivní) je správné zařazení do třídy „+“, *FP* (nesprávně pozitivní) je nesprávná klasifikace do třídy „+“(patří do „-“), *TN* (správně negativní) je správné zařazení do třídy „-“ a *FN* je nesprávná klasifikace do třídy „-“ (patří do „+“). BERKA (2003).

Tab. 6 Matice záměn, podle BERKA (2003).

	Klasifikace systémem	
Správné zařazení	+	–
+	*TP*	*FN*
–	*FP*	*TN*

Pro stanovení kvality získaných znalostí se používají charakteristiky, které z matice záměn vycházejí. Celková správnost (overal accuracy) *Acc* je relativní počet správných rozhodnutí klasifikátoru

$$Acc = \frac{TP + TN}{TP + TN + FP + FN}, \tag{5.10}$$

celková chyba (overal error) *Err* je relativní počet chybných rozhodnutí klasifikátoru

$$Err = \frac{FP + FN}{TP + TN + FP + FN}. \tag{5.11}$$

Celková chyba je komplementární k celkové správnosti, tudíž ji můžeme zapsat

$$Err = 1 - Acc. \tag{5.12}$$

Pokud jsou třídy v datech rozloženy nerovnoměrně, což je i případ hodnot cílového atributu spokojenost občanů s kvalitou okolního životního prostředí, bude celková správnost dávat zkreslený obraz o nalezených znalostech. V takovém případě je vhodnější sledovat správnost pro jednotlivé třídy

$$Acc_{+} = \frac{TP}{TP + FP}, \tag{5.13}$$

$$Acc_{-} = \frac{TN}{TN + FN}. \tag{5.14}$$

Přesnost a úplnost (precision and recall) jsou další užívané charakteristiky. Přesnost udává počet správně zařazených objektů dané třídy a úplnost počet správně zařazených objektů v dané třídě. Přesnost odpovídá správnosti pro třídu podle vztahu (5.13), úplnost vypočítáme

$$Úplnost = \frac{TP}{TP + FN}. \tag{5.15}$$

Senzitivita a specificita (sensitivity and specificity) jsou charakteristiky převzaté z medicíny. V případě nasazení léku nás zajímá, u kolika pacientů lék zabere (senzitivita), a zda lék zabírá pouze na danou chorobu (specificita). Senzitivita odpovídá úplnosti podle vztahu (5.15). Specificita se z matice záměn vypočítá

$$Specificita = \frac{TN}{TN + FP}. \tag{5.16}$$

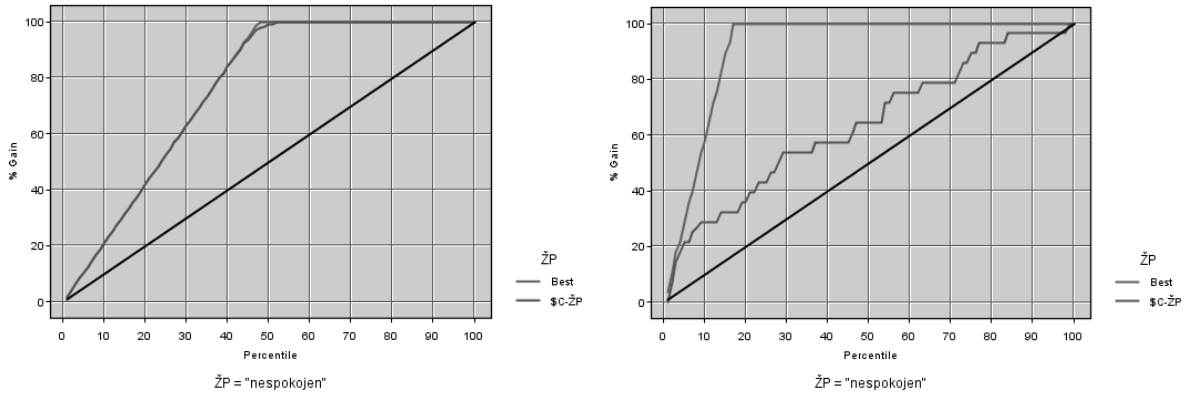
5.6.1 Vyhodnocení modelů rozhodovacích stromů

Rozdílný počet testovacích objektů u některých modelů je dán rozdělením dat pomocí generátoru náhodných čísel v Clementine. Tento počet neměl vliv na výslednou správnost klasifikace objektů. Následující matice záměn ukazují postupně výsledky klasifikace všech výše popsaných modelů.

Modely DT1, DT2 a DT3 (viz matice záměn - tabulky 7, 8 a 9) vykazují obdobné výsledky, které se vyznačují poměrně vysokou správností klasifikace u třídy „spokojen“, ale nízkou u třídy „nespokojen“. U modelu DT3 byl rozšířen interval hodnoty cílového atributu „nespokojen“ z [0, 5] na [0, 6], čímž se zvýšilo její zastoupení v souboru, ale výsledek se ještě zhoršil. Porovnání křivek navýšení pro méně zastoupené hodnoty „nespokojen“ modelu DT1 je znázorněno na obr. 21. Na levém grafu je „ideální“ správnost klasifikace dosažená na trénovacích datech, na pravém je správnost klasifikace dosažená na datech testovacích. Podstatný je samozřejmě výsledek na testovacích datech. Navýšení je v Clementine definováno jako poměr správných zařazení v každém kvantilu. Navýšení je počítáno jako (množství správných zařazení v kvantilu / všechna správná zařazení) × 100 %. Křivka, která je na každém grafu vždy vlevo, je nejlepší možná klasifikace, prostřední je křivka navýšení a na diagonále se nachází křivka odpovídající náhodnému výběru.

Tab. 7 Model DT1 - matice záměn

	Klasifikace systémem		Celková správnost
Správné zařazení	spokojen	nespokojen	
spokojen	112	27	
nespokojen	16	12	
Správnost třídy	87,50%	30,77%	74,3%



Obr. 21 Model DT1 - porovnání křivek navýšení

Tab. 8 Model DT2 - matice záměn

	Klasifikace systémem		Celková správnost
Správné zařazení	spokojen	nespokojen	
spokojen	119	20	
nespokojen	17	11	
Správnost třídy	87,50%	35,48%	77,8%

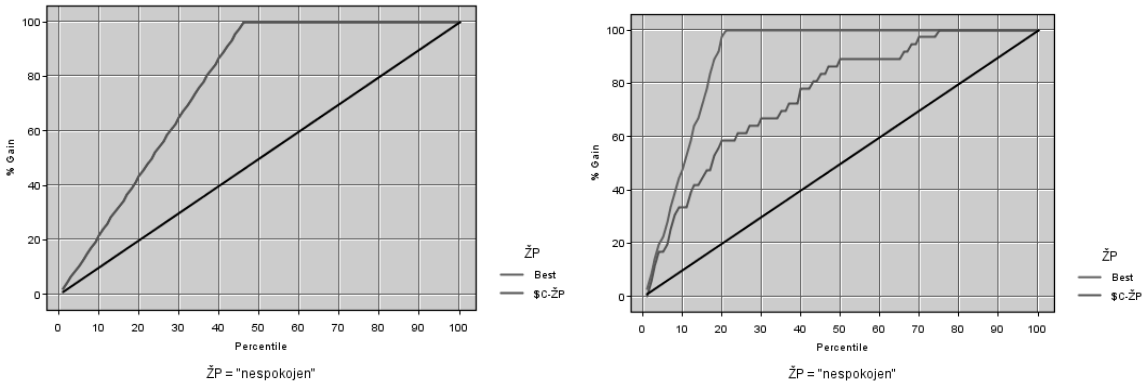
Tab. 9 Model DT3 - matice záměn

	Klasifikace systémem		Celková správnost
Správné zařazení	spokojen	nespokojen	
spokojen	103	36	
nespokojen	14	14	
Správnost třídy	88,03%	28,00%	70,1%

Model DT4 (viz matice záměn - tab. 10) dosahuje nejvyšší celkové správnosti z vytvořených modelů - 83 % a druhé nejvyšší správnosti klasifikace u třídy „nespokojen“. Důvodem, ale současně nevýhodou tohoto modelu je zřejmě využití všech atributů jako vstupních. Rozhodovací strom je díky vysokému počtu atributů obtížněji interpretovatelný. Porovnání křivek navýšení modelu DT4 je znázorněno na obr. 22. Na levém grafu je opět „ideální“ správnost klasifikace dosažená na trénovacích datech, na pravém je klasifikace dosažená na datech testovacích.

Tab. 10 Model DT4 - matice záměn

	Klasifikace systémem		Celková správnost
Správné zařazení	spokojen	nespokojen	
spokojen	132	10	
nespokojen	20	16	
Správnost třídy	86,84%	61,54%	83,1%



Obr. 22 Model DT4 - porovnání křivek navýšení

Modely DT5, DT6 a DT7 s algoritmy CHAID, QUEST a C&RT dosahují na těchto datech horší výsledky než algoritmus C5.0, celková správnost byla nejnižší a správnost klasifikace u třídy „nespokojen“ byla i přes zařazení všech atributů také nízká (viz matice záměn - tabulky 11, 12 a13).

Tab. 11 Model DT5 - matice záměn

	Klasifikace systémem		Celková správnost
Správné zařazení	spokojen	nespokojen	
spokojen	107	35	
nespokojen	19	17	
Správnost třídy	84,92%	32,69%	69,7%

Tab. 12 Model DT6 - matice záměn

	Klasifikace systémem		Celková správnost
Správné zařazení	spokojen	nespokojen	
spokojen	93	49	
nespokojen	11	25	
Správnost třídy	89,42%	33,78%	66,3%

Tab. 13 Model DT7 - matice záměn

	Klasifikace systémem		Celková správnost
Správné zařazení	spokojen	nespokojen	
spokojen	98	44	
nespokojen	16	20	
Správnost třídy	85,96%	31,25%	66,3%

5.6.2 Vyhodnocení fuzzy modelů

Modely FIS1 a FIS2 (viz matice záměn - tabulky 14 a 15) vykazují obdobné výsledky celkové správnosti jako rozhodovací stromy a vyznačují se vyšší správností klasifikace u třídy „nespokojen“. Z výsledků testů modelů typu FIS1 a FIS2 vyplývá, že lze zvyšovat úpravami parametrů funkcí příslušnosti a doplňováním báze pravidel správnost klasifikace u třídy „nespokojen“, ale vždy jen za cenu snížení správnosti klasifikace u třídy „spokojen“.

Tab. 14 Model FIS1 - matice záměn

	Klasifikace systémem		Celková správnost
Správné zařazení	spokojen	nespokojen	
spokojen	113	24	
nespokojen	34	27	
Správnost třídy	76,87%	52,94%	70,7%

Tab. 15 Model FIS2 - matice záměn

	Klasifikace systémem		Celková správnost
Správné zařazení	spokojen	nespokojen	
spokojen	118	19	
nespokojen	40	21	
Správnost třídy	74,68%	52,50%	70,2%

Model FIS3 (viz matice záměn - tab. 16) dosahuje nejvyšší celkové správnosti z vytvořených fuzzy modelů - 72 %. Ze všech modelů má nejvyšší správnost klasifikace u třídy „nespokojen“ - 71 %. Z výsledků testů modelů typu FIS3 vyplývá, že lze výrazně zvyšovat správnost klasifikace změnami typů funkcí příslušnosti a změnou defuzzifikační metody.

Tab. 16 Model FIS3 - matice záměn

	Klasifikace systémem		Celková správnost
Správné zařazení	spokojen	nespokojen	
spokojen	133	4	
nespokojen	51	10	
Správnost třídy	72,28%	71,43%	72,2%

Naopak modely FIS4 a FIS5 dosahují nejhorších výsledků ze všech vytvořených modelů jak v celkové správnosti, tak i správnosti klasifikace u obou tříd. Model FIS 4 (viz matice záměn - tab. 17) má bázi pravidel vytvořenu ze všech 501 generovaných pravidel. Ty sice mohou lépe pokrýt možné kombinace funkcí příslušnosti a posílit úplnost souboru pravidel, ale zároveň se vyskytují četné duplicity. Největším problémem tohoto způsobu naplňování báze pravidel je nekonzistence této báze (viz část 3.5.2). Existují zde pravidla, která mají stejný antecedent, ale různý konsekvent. Systém pak objekty s těmito antecedenty nedokáže zařadit a snižuje se správnost klasifikace.

Tab. 17 Model FIS4 - matice záměn

	Klasifikace systémem		Celková správnost
Správné zařazení	spokojen	nespokojen	
spokojen	109	28	
nespokojen	44	17	
Správnost třídy	71,24%	37,78%	63,3%

Model FIS5 sice dosahuje ve srovnání se zde uváděnými modely nejhorších výsledků (viz matice záměn - tab. 18), ale díky využití socio-demografických atributů by při optimalizaci mohl mít větší význam pro podporu rozhodování regionálního managementu, než modely užívající data získávaná náročným dotazníkovým šetřením.

Tab. 18 Model FIS5 - matice záměn

	Klasifikace systémem		Celková správnost
Správné zařazení	spokojen	nespokojen	
spokojen	91	46	
nespokojen	33	28	
Správnost třídy	73,39%	37,84%	60,1%

Na závěr části věnované testování a hodnocení rozhodovacích stromů a fuzzy modelů je uveden v tab. 19 přehled modelů a výsledků jejich celkové správnosti klasifikace a správnosti klasifikace pro obě třídy.

Tab. 19 Přehled výsledků klasifikace

	Správnost třídy		Celková správnost
	spokojen	nespokojen	
Model DT1	88%	31%	74%
Model DT2	88%	35%	78%
Model DT3	88%	28%	70%
Model DT4	**87%**	**62%**	**83%**
Model DT5	85%	33%	70%
Model DT6	89%	34%	66%
Model DT7	86%	31%	66%
Model FIS 1	77%	53%	71%
Model FIS 2	75%	53%	70%
Model FIS 3	**72%**	**71%**	**72%**
Model FIS 4	71%	38%	64%
Model FIS 5	73%	38%	60%

Z přehledu výsledků nevyplývá jednoznačná úspěšnost jedné z vybraných metod. Pomocí rozhodovacích stromů i fuzzy inferenčního systému typu Mamdani lze sestavovat pro řešení dané klasifikační úlohy modely ve srovnatelné kvalitě.

6 Závěr

Diplomová práce je zaměřena na problematiku hodnocení spokojenosti občanů se životem v obci a spokojenosti s různými rysy samosprávy. Popisuje význam informací získaných tímto hodnocením pro rozhodování orgánů veřejné správy, zejména regionálního managementu. Hlavním cílem diplomové práce je návrh metod pro klasifikaci občanů do tříd podle míry jejich spokojenosti. Byly navrženy a popsány dvě metody založené na rozhodovacích pravidlech - algoritmy pro tvorbu rozhodovacích stromů a fuzzy inferenční systémy. Algoritmy pro tvorbu rozhodovacích stromů jako metody strojového učení umožňují efektivní získávání znalostí z rozsáhlých datových souborů a jejich srozumitelnou interpretaci. Vzhledem k nepřesnosti pojmů užívaných při měření spokojenosti občanů jsou další vhodnou metodou fuzzy inferenční systémy, které umožňují tuto nepřesnost zahrnout.

Problematika regionálního managementu i problematika pravidlových systémů jsou v literatuře jako samostatné disciplíny velice dobře popsány. Jiná situace je však v oblasti aplikace pravidlových systémů na řešení problémů regionálního managementu. Metodika měření spokojenosti občanů sice existuje, ale další hledání informací se musí omezit již jen na hledání analogií, převážně s problémy řešenými v ekonomické oblasti. Sociologie má velice dobře propracované statistické metody a techniky sběru a analýzy dat a existují i práce týkající se aplikace fuzzy logiky v humanitních vědách, ale nalézt literaturu zaměřenou na aplikaci umělé inteligence v oblasti hodnocení spokojenosti občanů se mi nalézt nepodařilo.

Klasifikační úloha o spokojenosti občanů byla formalizována na základě formalizace obecné definice klasifikační úlohy a s pomocí analogií z jiných oborů. Úloha využívá již existujících výsledků dotazníkového šetření Spokojenost občanů s místní společenstvím (Obecná spokojenost občanů s různými rysy samosprávy). V roce 2004 byl proveden na území města první dotazníkový průzkum zaměřený na tento indikátor, který byl inspirací pro tuto práci. Tento indikátor umožňuje sběr srovnatelných údajů v rámci celé Evropy, v rámci srovnatelně velkých sídelních útvarů. Jako vstupní data jsou však využity „čerstvé“ výsledky šetření z podzimu roku 2007, vydané letos v lednu.

Vstupní data byla předzpracována v aplikaci MS Excel, k modelování spokojenosti občanů a vyhodnocení modelů je využito pro rozhodovací stromy prostředí SPSS Clementine a pro fuzzy inferenční systémy prostředí MATLAB. Základním problémem realizace klasifikačních metod je nalezení takových atributů, které dostatečně charakterizují jednotlivé objek-

ty, v tomto případě respondenty dotazníkového šetření. Ty pak umožňují správnou klasifikaci do tříd podle míry jejich spokojenosti s cílovým atributem. Jako cílový byl vybrán atribut vyjadřující spokojenost s kvalitou životního prostředí. Bylo vybráno několik rozdílných skupin vstupních atributů.

Po formalizaci klasifikační úlohy a stanovení atributů bylo navrženo několik skupin modelů pravidlových systémů klasifikace občanů podle jejich spokojenosti se životním prostředím v Chrudimi. Část skupin modelů využívá algoritmy pro tvorbu rozhodovacích stromů v Clementine a část jsou modely sestavené na základě fuzzy inferenčního systému typu Mamdani. U všech modelů byla testována celková správnost klasifikace a správnost klasifikace pro jednotlivé třídy.

Přestože většina vytvořených modelů nedosahovala příliš dobrých výsledků, testy prokázaly, že pro některé skupiny atributů lze funkční modely vytvářet. Hlavní cíl diplomové práce - návrh metod pro klasifikaci občanů do tříd podle míry jejich spokojenosti - lze proto považovat za splněný. Není ani dost dobře možné srovnávat správnost klasifikace modelů navrhovaných ve spolupráci s experty pro fyzikální nebo technické systémy. Technické systémy jsou ve srovnání se systémy sociálními lépe popsatelné a mají propracované metody simulace, které umožňují definovat vstupní atributy klasifikačních modelů. Velkým omezením při této práci byla i závislost na dodaném souboru dat, který se vyznačoval relativně vysokou mírou chybějících hodnot a výskytem chyb plynoucích z ne příliš šťastně zvolené stupnice hodnocení.

Modely navržené v této práci jsou prvními pokusy a k jejich další optimalizaci by byla nutná spolupráce s experty - sociology. Nabízí se možnosti využití např. hierarchických fuzzy inferenčních systémů, které umožňují výrazné snížení počtu pravidel, ale pro jejich smysluplné sestavení by byla spolupráce sociologa již nezbytná. Stejný předpoklad má i využití shlukové nebo fuzzy shlukové analýzy. Podařilo se sice sestavit poměrně úspěšné klasifikační modely, jejichž cílovým atributem byly shluky respondentů, ale interpretace těchto shluků je obtížná a nejednoznačná.

Některé modely byly aplikovány na data poskytnutá Městským úřadem Hodonín, která byla získána stejným dotazníkovým šetřením a mají obdobnou strukturu. Výsledky byly v porovnání s chrudimskými testovacími daty horší. Vzhledem k povaze dat a jejich zřejmé vazbě na daný region je přenositelnost jednotlivých souborů pravidel na jiná data problematická. Řešením je vytvoření nových modelů s využitím vlastních trénovacích dat příslušného regionu.

Původní záměr - sestavení „úspěšných“ modelů klasifikace občanů na základě vstupních socio-demografických dat, která jsou již obsažena v databázích, pro které není nutné získávat dodatečná data - se nepodařilo v plné míře naplnit. Velké možnosti však nabízí spolupráce při sestavování dotazníků pro další šetření. Na základě dosavadních výsledků lze tvrdit, že pokud se do dotazníků zařadí dostatečný počet otázek (atributů) týkajících se socio-demografických údajů o občanech, které jsou již v databázích a kterými disponuje regionální management, a současně otázek týkajících se hodnocení spokojenosti, vznikne soubor, který lze použít jako vhodná trénovací data pro klasifikační modely. Tyto modely, ať již na bázi rozhodovacích stromů nebo fuzzy inferenčních systémů, mohou dosahovat výsledků, které by byly přínosem pro rozhodování regionálního managementu.

7 Literatura

[1] ADAMČÍK, Stanislav. *Regionální politika a management regionů, obcí a měst*. 1. vyd. Ostrava: VŠB - Technická univerzita Ostrava., 2000. ISBN: 80-7078-837-0.

[2] BERKA, Petr. *Dobývání znalostí z databází*. 1. vyd. Praha: Academia, 2003. ISBN 80-200-1062-9.

[3] BRYCHTOVÁ, Šárka - DUPLINSKÝ, Josef. Image obce - pohled občana. In *Aktuální otázky rozvoje regionů*. 1. vyd. Pardubice: Univerzita Pardubice, 2005. s. 57 – 64. ISBN 80-7194-777-6.

[4] DRUCKER, Peter F. *Management: budoucnost začíná dnes.* 1. vyd. Praha: Management Press, 1992. ISBN 80-85603-00-4.

[5] DRUCKER, Peter F. - MACIARIELLO, Joseph A. *Drucker na každý den.* 1. vyd. Praha: Management Press, 2006. ISBN 80-7261-140-2.

[6] ČSN EN ISO 9000:2001, Systémy managementu jakosti – Základy, zásady a slovník. Praha: Český normalizační institut, c2002.

[7] DVOŘÁK, Jiří. *Expertní systémy* [online]. c2004 [cit. 2007-09-03]. Dostupné z: <http://www.uai.fme.vutbr.cz/~jdvorak/Opory/ExpertniSystemy.pdf>.

[8] ERISA. The European Regional Information Society Association. *Propagace ekosystémů a regionálního rozvoje na podporu regionálních operačních programů* [online]. 3.10.2007 [cit. 2008-04-17]. Dostupné z: <http://files.opaals.org/PEARDROP/PEARDROP%20Initial%20Guides/Czech>.

[9] EVROPSKÁ KOMISE. *Digital Business Ecosystems* [online]. c2007 [cit. 2007-10-11]. Dostupné z: <http://www.digital-ecosystems.org/book/de-book2007.html> ISBN 92-79-01817-5.

[10] FOTR, Jiří, et al. *Manažerské rozhodování: postupy, metody a nástroje*. 1. vyd. Praha: Ekopress, 2006. ISBN 80-86929-15-9.

[11] GIDDENS, Anthony. *Sociologie*. 1. vyd. Praha: Argo, 1999. ISBN: 80-7203-124-4.

[12] GROSPIČ, Jiří. Veřejný management regionálního rozvoje v obcích a krajích a některé právní a organizační otázky. In *Management regionální politiky a reforma veřejné správy*. Wokoun, R., Mates, P. (eds.). Praha: Linde, 2006. s. 11 - 36. ISBN: 80-7201-547-8.

[13] GRZYMALA-BUSSE, Jerzy - ZIARKO, Wojciech. Rough Sets and Data Mining. *Encyclopedia of Data Warehousing and Mining*. Ed. John Wang. Vol. 2. Hershey, Pennsylvania: Idea Group Reference, 2006. s. 973-977. 2 vols. *Gale Virtual Reference Library*. Gale. University of Pardubice. 17.4. 2008. <http://find.galegroup.com/gvrl/infomark.do?&contentSet=EBKS&type=retrieve&tabID=T002&prodId=GVRL&docId=CX3466900202&eisbn=1-59140-559-9&source=gale&userGroupName=cdv&version=1.0>.

[14] HEARST, Marti. *Marti Hearst: What Is Text Mining?* [online]. c2003, [cit. 2005-10-20]. Dostupné z: <http://www.sims.berkeley.edu/~hearst/text-mining.html>.

[15] CHAPMAN, Pete, et al. *CRISP-DM 1.0: Step-by-step data mining guide* [online]. 15.8.2000 [cit. 2008-04-09]. Dostupné z: <http://www.crisp-dm.org/CRISPWP-0800.pdf>.

[16] JALOVECKÁ, Martina. 5. konference kvality ve veřejné správě – Paříž 2008. *Veřejná správa: týdeník vlády České republiky*. 2008, roč. 11, č. 5, s. VII-VIII. ISSN: 1213-6581.

[17] JURA, Pavel. *Základy fuzzy logiky pro řízení a modelování*. 1. vyd. Brno: Vysoké učení technické v Brně, nakladatelství VUTIUM, 2003. ISBN: 80-214-2261-0

[18] KLIMEŠ, Lumír. *Slovník cizích slov*. 3. vyd. Praha: Státní pedagogické nakladatelství, 1985.

[19] KOORY, Jerry L. - MEDLEY, Don B. *Management information systems: planning and decision making*. Cincinnati, Ohio: South Western Publishing, 1987.

[20] KOŠČOVÁ, M. - EXNAR, Z. - KŘUPKA, J. *Systémy automatického riadenia*. Liptovský Mikuláš: Vojenská akadémia v Liptovskom Mikuláši, 1997. ISBN: 80-8040-053-9.

[21] KOTEK, Zdeněk. *Adaptivní a učící se systémy*. 1. vyd. Praha : Státní nakladatelství technické literatury, 1980.

[22] KOTEK, Zdeněk, et al. *Metody rozpoznávání a jejich aplikace*. 1. vyd. Praha: Academia, 1993. ISBN: 80-200-0297-9.

[23] KRATOCHVÍL, Ivan. O *řízení vážně i s úsměvem.* Kladno: Ing. Vladimír Macek, [2000?]. ISBN 80-86091-32-5.

[24] LUKÁŠ, Martin. *Městský informační management*. 1. vyd. Praha: Grada Publishing, 2000. ISBN 80-7169-554-8.

[25] MAŘÍK, Vladimír. Rozpoznávání. In *Umělá inteligence 1*. Mařík V., Štěpánková O., Lažanský J. a kol. 1. vyd. Praha: Academia, 1993. s. 145 – 167. ISBN 80-200-0496-3.

[26] MAŘÍK, Vladimír – ZDRÁHAL Zdeněk. Expertní systémy. In *Umělá inteligence 2*. Mařík V., Štěpánková O., Lažanský J. a kol. 1. vyd. Praha: Academia, 1997. s. 15 – 77. ISBN 80-200-0504-8.

[27] THE MATHWORKS. *Fuzzy Logic Toolbox: User's Guide.* Version 2 [online]. 7. 12. 1998. [cit. 2007-11-13]. Dostupné z: <http://som.yale.edu/unix/research_computing/pdf_doc/fuzzy/fuzzy_tb.pdf>.

[28] MINISTERSTVO VNITRA. *Kvalita ve veřejné správě* [online]. c2005 [cit. 2008-04-09]. Dostupné z: <http://www.mvcr.cz/odbor/moderniz/koncepce/kvalita.html>.

[29] NKÚ. *Prostředky státního rozpočtu vynaložené na Národní plán vyrovnávání příležitostí pro občany se zdravotním postižením* [online]. 13. 7. 2004 [cit. 2008-04-21]. Dostupné z: <http://www.nku.cz/scripts/rka/detail.asp?cisloakce=03/35&rok=0&sestava=0K>.

[30] NOVÁK, Josef. *Evropské indikátory udržitelného rozvoje v praxi měst České republiky* [online]. 25.09.2006, poslední revize 1.2.2007 [cit. 2007-09-02]. Dostupné z: <http://www.timur.cz/index2.php?option=com_docman&gid=13&task=doc_view&Itemid=38>.

[31] NOVÁK, Vilém. *Základy fuzzy modelování*. 1. vyd. Praha: BEN - technická literatura, 2000. ISBN: 80-7300-009-1.

[32] NOVOTNÝ, Jiří. *Stroje a množiny Zdzdislawa Pawlaka* [online]. 27. 9. 2006 [cit. 2007-01-09]. Dostupné z: <http://math.fce.vutbr.cz/~pribyl/workshop_2006/prispevky/Novotny.pdf>

[33] OLEJ, Vladimír. *Modelovanie ekonomických procesov na báze výpočtovej inteligencie.* 1. vyd. Hradec Králové: Miloš Vognar – M&V, [2003]. ISBN 80-90324-9-1.

[34] PPP Centrum. *PPP Centrum Stručně o PPP* [online]. [cit. 2008-04-18]. Dostupné z: <http://www.pppcentrum.cz/index.php?cmd=page&id=122>.

[35] PRACOVNÍ SKUPINA. Pracovní skupina pro měření, sledování a hodnocení udržitelnosti rozvoje na místní úrovni, Expertní skupina pro městské prostředí. *K trvale udržitelným místním profilům: Technická zpráva* [online]. c2000, [cit. 2007-07-01]. Dostupné z: <http://www.reccr.cz/download/indikator/technicka_zprava_cz.rtf>

[36] PŮČEK, Milan, et al. *Měření spokojenosti v organizacích veřejné správ -: soubor příkladů*. [online]. c2005, [cit. 2008-04-09].1.vyd. Praha: Ministerstvo vnitra České republiky, úsek veřejné správy, odbor modernizace veřejné správy, 2005. ISBN 80-239-6154-3. Dostupné z: <http://www.mvcr.cz/odbor/moderniz/spokojenost_final.pdf>.

[37] RECKLIES, Dagmar. *7-S-Model - Publications - Management Portal* [online]. c2001 [cit. 2008-04-18]. Dostupné z: <http://www.themanager.org/models/7s%20model.htm>.

[38] ROSS, Timothy J. *Fuzzy Logic with Engineering Applications*. 2nd ed. Chichester : John Wiley & Sons, 2004. ISBN: 0-470-86075-8

[39] RYCHLÝ, Marek. *Klasifikace a predikce* [online]. 11.1.2006 [cit. 2007-11-22]. Dostupné z: <http://www.fit.vutbr.cz/~rychly/docs/classification-and-prediction/classification-and-prediction.pdf>

[40] SHEN, Hong. Flexible Mining of Association Rules. *Encyclopedia of Data Warehousing and Mining.* Ed. John Wang. Vol. 1. Hershey, Pennsylvania: Idea Group Reference, 2006. s. 509-513. 2 vols. *Gale Virtual Reference Library*. Gale. Univerzita Pardubice. 17.4.2008 <http://find.galegroup.com/gvrl/infomark.do?&contentSet=EBKS&type=retrieve&tabID=T002&prodId=GVRL&docId=CX3466900106&eisbn=1-59140-559-9&source=gale&userGroupName=cdv&version=1.0>.

[41] SIEGL, Milan. *Management - obecné pojetí a veřejný sektor* [online]. 2002 [cit. 2008-03-20]. Dostupné z: <http://library.upce.cz/Sources/CL309.pdf>.

[42] SPSS. *SPSS Classification Trees* [online]. 5. 2. 2008 [cit. 2007-11-22]. Dostupné z: <http://www.spss.cz/sw_mcla.htm>.

[43] SUR. *Strategie udržitelného rozvoje České republiky* [online]. 21.12.2004 [cit. 2007-12-06]. Dostupné z: <http://www.env.cz/AIS/web-pub.nsf/$pid/MZPISF7Z6L7V>

[44] ŠTACH, Jaromír. *Základy teorie systémů.* Praha : Státní nakladatelství technické literatury, 1982.

[45] ŠVARC, Ivan. *Základy automatizace* [online]. 16.11.2002, poslední revize 9.12.2002 [cit. 2004-10-19]. Dostupné z: <http://www.fme.vutbr.cz/opory/pdf/ZakladyAutomatizace.pdf>.

[46] TIMUR. *Metodika dotazníkového průzkumu* [online]. c2006, [cit. 2008-04-09]. Dostupné z: <http://www.timur.cz/cz/dokumenty/metodicke-listy-eci/index.php?limit=5&limitstart=10&dir=DESC&order=name>.

[47] TUHÁRSKY, Peter - VOSKÁR, Štefan. Open-source technológie v samospráve mesta Banská Bystrica. In *Sborník konference ISSS 2008* [online]. c2008 [cit. 2008-05-05]. Dostupné z: <http://www.isss.cz/archiv/2008/download/isss2008.pdf>.

[48] TURBAN, Efraim - MEREDITH, Jack R. *Fundamentals of management science.* 5th ed. Homewood: Irwin, 1991. ISBN: 0-256-08373-8.

[49] TURBAN, E. – ARONSON, J. E. – LIANG, T. P.: *Decision Support Systems and Intelligent Systems.* 7th ed. Upper Saddle River: Pearson Education, Inc., 2005. ISBN 0-13-046106-7.

[50] VORLÍČEK, Zdeněk. Strategie Národní politiky kvality v České republice na období let 2008 až 2013 pro vyšší kvalitu života občanů České republiky. *Veřejná správa: týdeník vlády České republiky*. 2008, roč. 11, č. 5, s. III-V. ISSN: 1213-6581.

[51] VESELÁ, Jana. *Sociologický výzkum a jeho metody.* 1. vyd. Pardubice: Univerzita Pardubice, 2002. ISBN 80-7194-466-1.

[52] WRIGHT, Glen. Úvod do managementu veřejné správy. In *Management veřejné správy: teorie a praxe: zkušenosti z transformace veřejné správy ze zemí střední a východní Evropy*. Wright, G., Nemec, J. (eds.). 1. vyd. Praha: Ekopress, 2003. s. 32 - 39. ISBN 80-86119-70-X.

[53] ZADEH, Lotfi A. Forward. In *Fuzzy Logic Toolbox: User's Guide.* Version 2 [online]. 7. 12. 1998. [cit. 2007-11-13]. Dostupné z: <http://som.yale.edu/unix/research_computing/pdf_doc/fuzzy/fuzzy_tb.pdf>.

[54] Zákon č. 1/1993 Sb., Ústava České republiky. [online]. [cit. 2008-04-18]. Dostupné z: <http://www.zakonycr.cz/seznamy/001-1993-Sb-ustava-ceske-republiky.html>.

[55] Zákon č. 128/2000 Sb., o obcích (obecní zřízení). [online]. [cit. 2008-04-18]. Dostupné z: <http://www.sagit.cz/pages/sbirkatxt.asp?zdroj=sb00128&cd=76&typ=r>.

[56] Zákon č. 129/2000 Sb., o krajích (krajské zřízení). [online]. [cit. 2008-04-18]. Dostupné z:<http://www.sagit.cz/pages/sbirkatxt.asp?zdroj=sb00129&cd=76&typ=r>.

[57] ŽELEZNÝ, F. - KLÉMA, J. - ŠTĚPÁNKOVÁ, O. Strojové učení v dobývání znalostí. In *Umělá inteligence 4*. Mařík V., Štěpánková O., Lažanský J. a kol. 1. vyd. Praha: Academia, 2003. s. 355 – 406. ISBN 80-200-1044-0.

Příloha 1: Dotazník – Indikátor 1

Dotazník – Indikátor 1

Níže uvedené údaje vyplní pouze JEDNA osoba (označená jako příjemce)

Obecné informace
a) **Pohlaví:** ☐ muž ☐ žena
b) Věk _____
c) Zaměstnání ☐ student ☐ zaměstnaný ☐ nezaměstnaný ☐ důchodce
d) adresa bydliště(městská část) : _____
e) v případě, že kouříte, kolik vykouříte cigaret denně _____-

1.a) Jak jste spokojen(a) s vaší obcí, jako s místem, kde žijete a pracujete? *(Vyberte prosím pouze jednu možnost)*
☐ velmi spokojen
☐ mírně spokojen
☐ mírně nespokojen
☐ velmi nespokojen

1.b) Vyjádřete úroveň vaši spokojenosti na procentní stupnici (0 velice nízká, 10 velmi vysoká)
Skóre: _____
Prosím přiřaďte ke každé otázce bodové skóre (podobně jako v předchozí otázce 0 velice nízká spokojenost, 10 velmi vysoká)
2) Jak jste spokojen s mezilidskými vztahy ve vaší obci?_____
3) Jak jste spokojen s možnostmi provozovat své záliby a koníčky? _____
4) Jak jste spokojen se základními veřejnými službami (zdravotní a sociální služby, školy, veřejná doprava, atd.)? _____
Z toho jak jste spokojen se :
a) zdravotními službami ? _____
b) sociálními službami ?_____
c) mateřskými a základními školami ? _____
d) středními školami ? _____
e) městskou hromadnou dopravou ? _____
f) fungováním Technických služeb ? _____
g) fungováním Městské policie ? _____
5) Jak jste spokojen s fungováním Městského úřadu ? _____
Z toho jak jste spokojen se :
a) jak se cítíte informován/s ze strany Městského úřadu ? _____
b) jak hodnotíte internetové stránky města ? (www.chrudim-city.cz) _____
c) jak hodnotíte obsahovou náplň Chrudimského zpravodaje ? _____
d) jaký pociťujete vliv činnosti Městského úřadu na život ve městě ? _____
e) je pro Vás vyjadřování úředníků při jednáních srozumitelné ? _____
f) jak hodnotíte zavedení vyvolávacího a objednávacího systému ? _____
6) Jak jste spokojen s kvalitou okolního životního prostředí? _____
7) Jak jste spokojen s možnostmi zaměstnání ve vašem městě? _____
8) Jak jste spokojen s možnostmi účastnit se místního plánování, rozhodování či kulatých stolů?_____
9) Jste dlouhodobě léčen/a nebo sledován/a lékařem pro nějaké onemocnění ? _____
☐ **Ano** **ne** **bez odpovědi**
10) Kolikrát za posledních 12 měsíců jste v Chrudimi navštívil zdravotnické zařízení ?
☐ **0** **1- 5** **6 – 10** **11 a více**
11) Kolik hodin v průměru se týdně věnujete aktivnímu pohybu nebo sportu ?
☐ **0** **1 – 3** **4 – 7** **8 – 10** **11 – 15** **16 – 20** **více než 20**
12) Domníváte se, že prostředí a životní podmínky ve vašem městě ovlivňují Vaše zdraví ?
☐ **Ano, velmi** **spíše ano** **nevím** **spíše ne** **určitě ne**

13) Zhodnoťte úroveň spokojenosti s vybranými oblastmi :
a) se svou bytovou situací _____
b) s možnosti třídit komunální odpad _____
c) s bezpečností ve městě ve dne _____
d) s možnosti relaxace a odpočinku _____

e) s nabídkou zdravé výživy v prodejnách, restauracích apod. _____
f) s bezpečností s ohledem na silniční dopravu, když jdete po městě pěšky _____
g) s možností ovlivňovat a podílet se na společenském dění ve městě _____
h) s možností jezdit ve městě bezpečně na kole_____
ch) se svým zdravotním stavem _____

Nyní prosím seřaďte dané oblasti/položky v závislosti na vašem osobním hodnocení důležitosti (1= nejvíce důležitý, 6 = nejméně důležitý)

Pořadí

_____ mezilidské vztahy
_____ možnost praktikovat vaše zájmy a záliby
_____ základní veřejné služby
_____ kvalita okolního životního prostředí
_____ možnost zaměstnání
_____ možnost účastnit se lokálního plánování (místní Agenda 21 apod.)

Přiřaďte prosím hodnocení na škále mezi 0 (nejméně) a 10 (nejvíce) každé položce v seznamu otázky a):

b) Jak bezpečné je podle vás: (*)

_____ Být přes den doma s nezamknutými dveřmi? ……..
_____ Nechat přes noc otevřené okno? ……..
_____ Chodit v noci po hlavních ulicích? ……..
_____ Chodit v noci po veřejných prostranstvích? ……..

c) Zhodnoťte prosím kvalitu následujících služeb:

_____ Sportovní zařízení ……..
_____ Divadlo a kino ……..
_____ Muzea a výstavní síně ……..
_____ Kulturní akce (Obžínky, jarmarky, Volba královny, plesy…)
_____ Knihovnu ……..

d) Jak dostupné jsou následující základní služby

_____ Praktičtí lékaři ……..
_____ Nemocnice ……..
_____ Sociální služby ……..
_____ Sociální bydlení ……..
_____ Policie ……..
_____ Školy ……..
_____ Veřejná doprava ……..

e) Zhodnoťte kvalitu následujících položek:

_____ Veřejné parky a zahrady, zeleň obecně ……..
_____ Zastavěné prostory (zastavěné plochy, výstavba ve městě)
_____ Odvoz odpadů a čištění ulic ……..
_____ Kvalita ovzduší ……..
_____ Hlučnost v noci ……..
_____ Hlučnost ve dne ……..

f) Jaký je váš názor na následující možnosti:

_____ Možnost odborné rekvalifikace ……..
_____ Podpora novým podnikatelským záměrům ……..
_____ Úroveň nezaměstnanosti ve vaší obci ……..
_____ Hospodaření s majetkem obce
_____ zákaz kouření v restauracích ……
_____zákaz kouření v nádražních prostorech ……..
_____zákaz kouření na úřadech ……
_____zákaz kouření ve školách ……
_____zákaz kouření na ostatních veřejných prostranstvích …….
_____

g) Zhodnoťte prosím efektivitu následujících možností zapojení do rozhodování ve vaší obci:

_____ Podílet se na místním plánovacím procesu (příprava územního plánu, místní Agenda 21 apod.)
_____ Stát se členem místní nevládní organizace ……..
_____ Podávat přímé žádosti/dotazy na místní úřad ……..
_____ Účast v komunálních volbách/referendech ……

Příloha 2: Metodika dotazníkového průzkumu

Podrobná metodika dotazníkového průzkumu pro naplnění indikátoru A.1 a A.3

Z registru obyvatel se vybere reprezentativní vzorek obyvatel následujícím způsobem:

- Odfiltrují se plnoletí obyvatelé (nad 18 let)
- Podle velikosti města se z registru náhodným výběrem vybere dostatečný počet respondentů (A) (pod 20 tisíc – **700** / 20 tisíc až 100 tisíc **850** / nad 100 tisíc – **1000** respondentů).
- *Náhodný výběr*: každý prvek (respondent) musí mít stejnou pravděpodobnost, že bude vybrán.
- Výběr z registru tedy musí být prováděn náhodným výběrem - vybere se každý desátý, dvacátý, … člen a vybírá se tak dlouho, dokud nedostaneme požadovaný počet respondentů (A).
- Při výběru počtu respondentů musíme počítat s dvěmi skutečnostmi:
  1. Nesmí být ve výběru dva nebo více členů jedné rodiny
  2. Někteří respondenti při dotazování nebudou moci (nebudou chtít) odpovídat.
- Z toho důvodu je nutné mít v zásobě další sadu respondentů (B), z které budeme nahrazovat původní soubor (tuto sadu vybíráme současně s tou první (A))

- Vybraný soubor respondentů (A) zkontrolujeme, zda neobsahuje dva nebo více členů z téže domácnosti. Pokud takové osoby výběr (A) obsahuje, přidáme stejné množství členů ze souboru (B), abychom dostali požadovaný počet respondentů.

Rozdělení vzorku respondentů mezi tazatele:
Každý tazatel obdrží seznam respondentů (seznam již neobsahuje osoby ze stejné domácnosti), se kterými by měl provézt vlastní dotazníkové šetření. Pokud bude deset tazatelů a bude potřeba provézt 850 rozhovorů, dostane každý tazatel 85 adres respondentů K těmto 85 adresám se přidá ještě 50% náhradníků se seznamu (B) – tedy pro každého tazatele 43 adres.

Tazatel při vlastním dotazování bude navštěvovat jednotlivé osoby z původního seznamu. Pokud nějakou osobu nedostihne doma na druhý pokus nebo pokud osoba odmítne odpovídat, nahradí tuto osobu členem ze seznamu náhradníků (respondenty ze seznamu náhradníků bere tazatel osoby popořadě – nikoliv podle známosti či blízkosti lokality – aby se nenarušil princip náhodnosti).

Tazatel skončí dotazování ve dvou případech:

1. Pokud má již provedl 85 rozhovorů (a zbylo mu několik osob ze seznamu náhradníků, se kterými dotazování neprováděl)
2. Pokud již vyčerpal s seznam náhradníků a nemá požadovaný počet 85 rozhovorů

V případě 2.si dojede pro další seznam náhradníků a z něho se pokusí udělat zbývající počet rozhovorů.

Tazatelé provádějí dotazování přímým rozhovorem: kladou otázky a zapisují odpovědi. V žádném případě nedávají dotazník z ruky k vyplnění respondentem samým.

TIMUR (2006)