

MODELOVÁNÍ SPOKOJENOSTI OBČANŮ VE VZTAHU K REGIONÁLNÍMU ROZVOJI A KVALITĚ ŽIVOTA

Miloslava Kašparová, Jiří Křupka, Jaromír Pírko

Univerzita Pardubice, Fakulta ekonomicko-správní, Ústav systémového inženýrství a informatiky

Abstract: *This paper presents a possibility of the decision trees theory using for a modelling of citizen satisfaction. In this paper the real data set of the citizen satisfaction with city Chrudim was analysed, pre-processed and used in the model.*

Keywords: *Regional development, regional management, method, model, classification systems, decision trees, citizen satisfaction*

1. Úvod

Regionální management je možné definovat 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.“ [1]. Tato definice je v podstatě analogií k definicím podnikového managementu a je zaměřena na ekonomickou stránku řešeného problému. V [1] je uvedeno její zobecnění: „... cílem regionálního managementu je prospěch a rozvoj regionu, uspokojování zájmů a potřeb obyvatelstva a jeho skupin a veřejný zájem“, které je v souladu s definicí v zákonu č. 128/2000 Sb., o obcích (obecní zřízení) a zákonu č. 129/2000 Sb., o krajích (krajské zřízení).

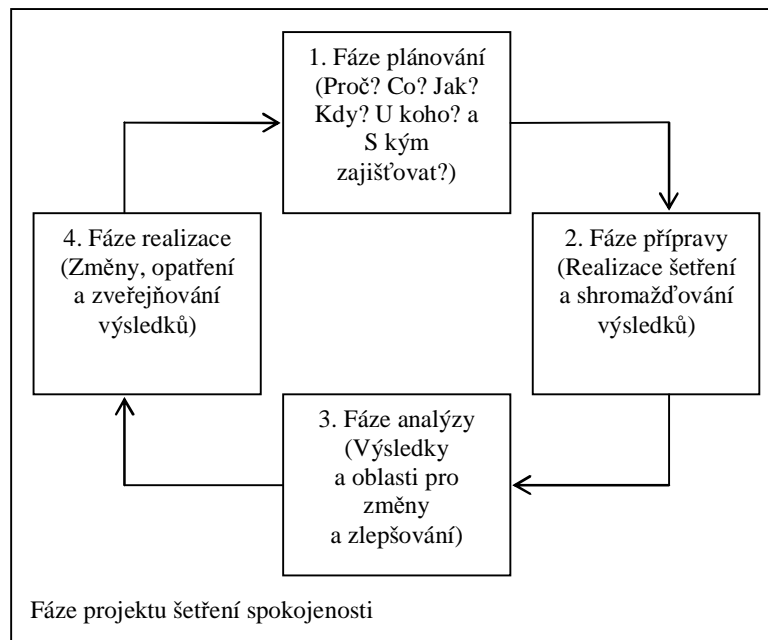
Ministerstvo vnitra podporuje zavádění nástrojů řízení kvality ve veřejné správě mj. i prostřednictvím metodického vedení a koordinace. Rada České republiky (ČR) pro jakost zpracovala návrh nové Strategie Národní politiky kvality v ČR na období let 2008 až 2013. Návrh Strategie vychází z analýzy výsledků dosavadního plnění Národní politiky podpory jakosti a na základě vyhodnocení současné situace definuje na příští období poslání, vizi, rámec a dlouhodobý strategický cíl. Tímto cílem 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 ČR je veden cestou udržitelného rozvoje [23]. Strategie má být přínosem pro podnikatelskou sféru, veřejnou správu i pro celou občanskou společnost. V oblasti veřejného sektoru má Strategie podporovat rozvoj kvality veřejné správy¹.

Mezi nejčastěji používané nástroje, metody a modely pro hodnocení kvality managementu veřejné správy mimo jiné patří například: 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, model CAF (Common Assessment Framework), EFQM (European Foundation for Quality Management), Benchmarking (porovnání výkonnosti), PDCA (Plan Do Check Act) [10,14]. V dané oblasti je možné použít také přímé měření spokojenosti občanů, vycházející z dotazníkového šetření.

¹ 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 a v říjnu 2008 v Paříži 5. konference kvality. Hlavním tématem konference se stává „Občan v srdci kvalitní veřejné správy“. Organizátoři popisují téma konference takto [7]: „Hlavním cílem veřejné správy je zlepšení života občanů. Abychom k tomuto cíli dospěli, je nezbytné zapojit do veřejné správy řízení procesů kvality.“

Způsob zjišťování informace pomocí přímého měření spokojenosti občanů poskytuje managementu zpětnou informaci o názorech občanů přímo od nich samotných. Přímé měření má své nesporné výhody, ale přináší i některé problémy. Ačkoliv je sám pojem „spokojenost“ definován normou ČSN EN ISO 9000:2001, je zřejmá² jeho nejednoznačnost.

Přes 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 - vychází z obecného cyklu PDCA (obr.1) a je upraven podle [14]:



Obr.1: Projekt a cyklus šetření spokojenosti

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 atd.). Na základě konzultací s pracovníky

² Velmi podobný problém, v uvedeném případě s pojmem „image“ při výzkumu „Image obce - pohled občana“, popisují [3]. 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 [12].

městských úřadů z několika českých měst vyplynulo [12], že některá města dříve šetření spokojenosti ve vybraných letech prováděla³.

2. Případová studie

Případová studie se týká modelování spokojenosti občanů s kvalitou životního prostředí ve městě Chrudim. Toto město využívá výsledků dotazníkových průzkumů ze sady Společné evropské indikátory jako zpětné vazby při řízení a rozvoji města. Dávají se na vědomí radě a zastupitelstvu města, odborům městského úřadu, organizacím města (Technické služby) a jsou zveřejňovány. Výsledky jsou využívány např. při rozhodování o přidělování dotací.

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 mining, DM)⁴. Podkladem pro modelování budou vybraná data získaná dotazníkovým šetřením Indikátoru 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 rozvoje⁵). Umožňuje sběr srovnatelných údajů v rámci celé Evropy a v rámci srovnatelně velkých sídelních útvarů. Projekt vychází z iniciativy skupin kolem Evropské komise [11,13]. Daný indikátor zjišťuje a vyčísluje subjektivní pocit spokojenosti občanů s městem, ve kterém žijí a pracují, a dílčí aspekty této spokojenosti.

První fází dobývání znalostí na základě metodiky CRISP-DM je pochopení cílů úlohy z manažerského hlediska a její převod na úlohu dobývání znalostí. V tomto případě, kde se zabýváme spokojeností občana s životním prostředím, je v manažerské roli regionální management. Manažerskou úlohou je hledání souvislostí mezi demografickými a jinými údaji o občanech, dílčími aspekty spokojenosti občanů, a především modelování spokojenosti občanů s životním prostředím s ohledem na regionální rozvoj a kvalitu života jedinců. Pokud by se podařilo definovat skupiny občanů, kteří jsou nespokojeni, mohou získané znalosti sloužit jako podklady pro stanovení priorit při dalším rozhodování a řízení rozvoje regionu. Model hodnocení spokojenosti občanů jako subsystém systému řízení [8,17,20,22] je znázorněn na obr.2.

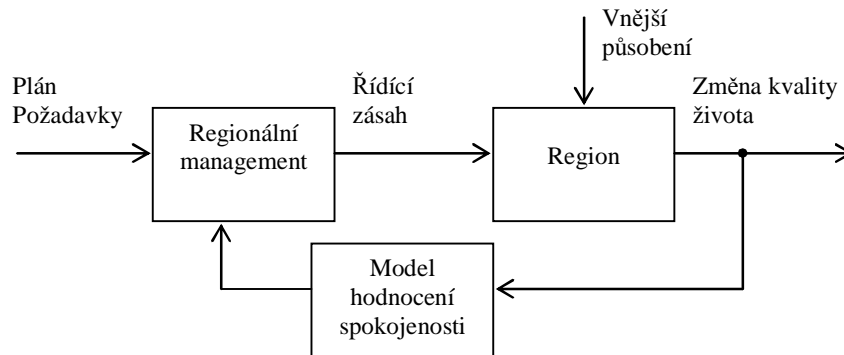
Z hlediska dobývání znalostí lze formulovat v této oblasti řada úloh. S ohledem na zaměření se jedná o nalezení takových demografických či jiných atributů o občanech, které mají vliv na určení spokojenosti občanů s životním prostředím, popřípadě na základě určení spokojenosti občana vytvořit klasifikátor, který na základě dostupných a použitých atributů zařadí každého občana podle uvedené spokojenosti do příslušné kategorie. Úloha proto může

³ 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.

⁴ Východí metodikou DM CRISP-DM člení celý proces data miningu do šesti základních etap. Jsou to [6]: porozumění problematice, porozumění datům, příprava dat, modelování, hodnocení a využití výsledků.

⁵ Udržitelný rozvoj je novým rámcem strategie civilizace 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 [19].

být pojata jako klasifikační, v které je cílem zařazení občanů do tříd podle určení jejich spokojenosti a v které bude využito rozhodovacích stromů.



Obr.2: Model hodnocení spokojenosti občanů jako zpětná vazba v systému řízení

2.1 Sběr a příprava dat

Dotazníkový průzkum byl proveden na území města Chrudim na podzim roku 2007. Na základě vybraných atributů vycházejících z indikátoru A1 byl vytvořen datový soubor, obsahující 701 záznamů (objektů, příkladů, respondentů) a 10 atributů (vlastností, proměnných, dotazníkových otázek). Přehled těchto atributů, které lze rozdělit na atributy obecné a specifické, je následující: obecné (*pohlaví, věk, zaměstnání*), specifické (*počet vykouřených cigaret denně, spokojenost se základními veřejnými službami (zdravotní a sociální služby, školy, veřejná doprava atd.), spokojenost s kvalitou okolního životního prostředí, počet hodin v průměru týdně aktivního pohybu nebo sportu, vyjádření, zda prostředí a životní podmínky ve městě mají vliv na zdraví dotazovaného, úroveň spokojenosti s možností relaxace a odpočinku*). Vzhledem ke skutečnosti, že dané atributy byly hodnoceny různě [12] (tab. 1), bylo nutné atributy i jejich hodnoty upravit do formy použitelné pro zvolený algoritmus.

Tab. 1: Přehled původních 10 atributů

Popis atributu	Měřitelnost
ID respondenta	Číslo
Pohlaví	Muž/žena
Věk	Číslo
Zaměstnání	Student (1) zaměstnaný (2) nezaměstnaný (3) důchodce (4)
Počet cigaret denně	Číslo
Spokojenost s veřejnými službami	Bodové hodnocení v rozsahu 0 -10, kde 0 je velice nízká, 10 je velmi vysoká)
Počet hodin v průměru týdně aktivního pohybu nebo sportu	Hodnocení 0, popřípadě výběr z intervalu <1,3>, <4,7>, <8,10>, <11,15>, <16,20> a více než 20
Vyjádření, zda prostředí a životní podmínky ve městě mají vliv na zdraví dotazovaného	Pomocí stupnice v rozsahu 1 – 5, kde 1 je ano velmi, 2 je spíše ano, 3 je nevím, 4 je spíše ne, 5 je ne
Úroveň spokojenosti s možností relaxace a odpočinku	Bodové hodnocení v rozsahu 0 -10, kde 0 je velice nízká, 10 je velmi vysoká)

Spokojenost s kvalitou okolního životního prostředí	Bodové hodnocení v rozsahu 0 -10, kde 0 je velice nízká, 10 je velmi vysoká)
--	--

Na základě úprav souboru, tj. odstranění odlehlých a chybných hodnot, úpravy atributů (převod atributů na dichotomické proměnné a diskretizace proměnných *věk*, *počet cigaret denně* a proměnných hodnocených pomocí bodového hodnocení v rozsahu od 0 do 10 viz tab. 1) byl získán datový soubor obsahující 691 záznamů popsaných 33 atributy uvedenými v tab. 2.

Tab. 2: Atributy

Atribut	Popis atributu	Rozsah hodnot	Počet
a1	x4-student	{0,1}	691
a2	x4-zam	{0,1}	691
a3	x4-nezam	{0,1}	691
a4	x4- Duch	{0,1}	691
a5	x16-ano_velmi	{0,1}	691
a6	x16_spise_ano	{0,1}	691
a7	x16-nevim	{0,1}	691
a8	x16-spise_ne	{0,1}	691
a9	x16-ne	{0,1}	691
a10	x15-_sport_0	{0,1}	691
a11	x15-sport_1-3	{0,1}	691
a12	x15-_sport_4-7	{0,1}	691
a13	x15-_sport_8-10	{0,1}	691
a14	x15-_sport_11-15	{0,1}	691
a15	x15-_sport_16-20	{0,1}	691
a16	x15-_sport_vicnez20	{0,1}	691
a17	x3-vek_skup_29	{0,1}	691
a18	x3-vek_skup_41	{0,1}	691
a19	x3-vek_skup_50	{0,1}	691
a20	x3-vek_skup_82	{0,1}	691
a21	x6-nekurak	{0,1}	691
a22	x6_kurak_9	{0,1}	691
a23	x6_kurak_19	{0,1}	691
a24	x6_kurak_40	{0,1}	691
a25	x2_zena	{0,1}	691
a26	x2_muz	{0,1}	691
a27	x17_relax_nizka	{0,1}	691
a28	x17_relax_prumer	{0,1}	691
a29	x17_relax_vysoka	{0,1}	691
a30	x11_sluzby_nizka	{0,1}	691
a31	x11-sluzby_prumer	{0,1}	691
a32	x11_sluzby_vysoka	{0,1}	691
a33	x13-kvalitaZP	{1,2,3}	691

U atributu *počet cigaret denně* byla provedena diskretizace na předem zadaný počet intervalů, a to: hodnota 2 pro interval <1, 9>, hodnota 3 pro interval <10, 19>, hodnota 4 pro 20 cigaret denně a více. V případě nekuřáků byla přidělena hodnota 1. Atribut *věk* byl diskretizován na předem zadaný počet ekvifrekvenčních intervalů, a to: hodnota 1 pro

respondenty ve věku 15 až 29, respondenti ve věku od 30 do 41 let byli označeni hodnotou 2, hodnota 3 odpovídá intervalu <42, 50> a hodnota 4 respondentům ve věku od 51 do 82 let. V každém intervalu uvedeného atributu bylo zařazeno cca 175 objektů. U proměnné *počet hodin v průměru týdně aktivního pohybu nebo sportu* bylo provedeno ohodnocení jednotlivých intervalů hodnotami 1 nebo 0. V případě platnosti vybraného intervalu z uvedených (viz tab. 1) byl atribut ohodnocen číslem 1, jinak 0. Stejným způsobem byl zpracován i atribut *vyjádření, zda prostředí a životní podmínky ve městě mají vliv na zdraví dotazovaného*. U atributů hodnocených pomocí bodů v rozsahu od 0 do 10 bylo provedeno následující: body v intervalu od 0 do 3 byly označeny číslem 1 (spokojenost nízká), od 4 do 6 číslem 2 (spokojenost průměrná) a body od 7 do 10 (spokojenost vysoká). Jednalo se o proměnnou *spokojenost s veřejnými službami, úroveň spokojenosti s možností relaxace a odpočinku, spokojenost s kvalitou okolního životního prostředí*.

Data jsou uložena v matici \mathbf{M} , která je tvořena n řádky a m sloupci, kde $n = 691$ a $m = 33$:

$$\mathbf{M} = \begin{bmatrix} a_{11} & a_{12} & \mathbf{K} & a_{1m} \\ a_{21} & a_{22} & \mathbf{K} & a_{2m} \\ \mathbf{M} & \mathbf{M} & & \mathbf{M} \\ a_{n1} & a_{n2} & \mathbf{K} & a_{nm} \end{bmatrix}.$$

Řádky matice reprezentující sledované objekty lze pro i -tý objekt zapsat jako:

$$\mathbf{a}_i = [a_{i1}, a_{i2}, \mathbf{K}, a_{im}], \text{ kde } i = 1, 2, \dots, 691.$$

Data získaná při šetření spokojenosti občanů lze zpracovávat pomocí řady metod umělé a výpočetní inteligence. Pro řešení dané úlohy jsou zvoleny klasifikační systémy založené na pravidlech. Jedná se o rozhodovací (klasifikační) stromy.

2.2 Návrh klasifikátoru

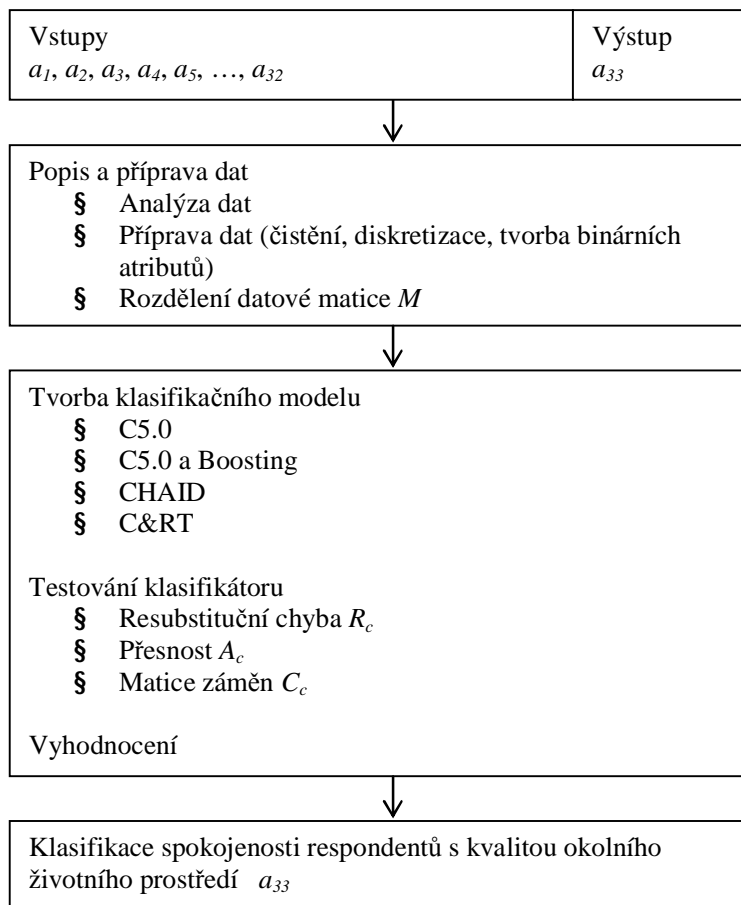
K nejznámějším klasifikačním modelům patří algoritmy pro vytváření rozhodovacích stromů. Rozhodovací strom [16] můžeme definovat jako strom (stromový graf), kde každý nelistový uzel stromu představuje test hodnoty atributu a větve vedoucí z tohoto uzlu možné výsledky testu. Listové uzly stromu jsou ohodnoceny 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. 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. Mezi nejznámější a běžně užívané algoritmy pro vytváření rozhodovacích stromů [2,18,25] patří například: C5.0, C4.5, C&RT (Classification & Regression Trees).

Modelování [21] je velmi rozšířenou pracovní poznávací metodou nacházející uplatnění v řadě oblastí společenské praxe. V našem případě se jedná o modelování spokojenost občanů s okolním životním prostředím na základě atributů uvedených v kap. 2.1. Jednotlivé

objekty jsou popsány pomocí 33 nezávislých atributů a 1 závislým (*spokojenost respondentů s kvalitou okolního životního prostředí*).

Pro vytvoření modelů a jeho ověření je nutné rozdělit datovou matici na matici, obsahující data trénovací a matici určenou k testování. Trénovací data jsou uložena v matici M_{TR} , tvořené n řádky a m sloupci, kde n odpovídá náhodnému výběru 66,67 % trénovacích příkladů a m odpovídá počtu vybraných atributů a testovací data v matici M_{TE} , která odpovídá 33,33 % výchozí datové matice. Pro modelování jsou zvoleny metody pro tvorbu rozhodovacích stromů (mezi známé algoritmy patří např. C&RT, C4.5, C5.0, Chi-square Automatic Interaction Detection (CHAID)). C4.5 je rozšířením verze ID3, umožňuje práci s numerickými atributy, chybějícími hodnotami, převod na pravidla i přežezávání. Verze C5.0 je rozšířením uvedené verze. C&RT umožňuje vytvářet vedle klasifikačních stromů i stromy regresní. Vytvořené stromy jsou binární. Algoritmus CHAID používá jako kritérium pro větvení χ^2 . Tento algoritmus seskupuje hodnoty kategoriálních atributů, při větvení se nevytváří tolik větví, kolik má atribut hodnot. Hodnoty atributu se postupně seskupují z původního počtu až do dvou skupin a poté se vybere atribut a jeho kategorizace, která je v daném kroku pro větvení nejlepší [2], více např. v [4,5,9,15,24]. Obecné schéma modelu uvádí obr.3.



Obr.3: Návrh klasifikačního modelu

2.3 Analýza výsledků

Cílem testování je určit, v kolika případech se klasifikátor shoduje s učitelem a v kolika se dopustil chyby. V daném případě lze využít matici záměn C_c , celkovou přesnost A_c popřípadě

celkovou chybu E_c [2, 5, 24]. Matice záměn C_c sleduje počty správně a nesprávně zařazených příkladů. Celková správnost A_c patří mezi jednoduché charakteristiky určující, jak jsou nalezené znalosti kvalitní. Výpočet celkové správnosti je následující:

$$A_c = (TP + TN)/(TP + TN + FP + FN), \quad (3)$$

kde TP znamená správně pozitivní, TN znamená správně negativní, FP znamená falešně pozitivní, FN je falešně negativní.

Z charakteristiky A_c lze získat celkovou chybu E_c , která se vypočte jako relativní počet chybných rozhodnutí systému:

$$E_c = (FP + FN)/(TP + TN + FP + FN), \quad (4)$$

V případě, že je zjišťován pouze počet správných či chybných rozhodnutí klasifikátoru, pak:

$$E_c = 1 - A_c \quad (5)$$

V případě, že provádíme testování na trénovacích datech, je možné získat také resubstituční chybu R_c [5, 24]. Tento způsob testování má však malou vypovídací schopnost z hlediska nalezených znalostí použitelných pro klasifikování nových případů. Přesto, že to není vhodný ukazatel správnosti klasifikace, je užitečné jej znát [24].

Výsledky dosažené pomocí navrženého klasifikátoru vytvořeného na základě algoritmu C5.0, CHAID a C&RT získané na testovacích datech (33,3 % respondentů z datové matice M) jsou uvedeny v tab. 3, tab. 4 a tab. 5.

V rámci algoritmu C5.0 byl využit i tzv. boosting, kdy vznikají postupně modely s rostoucí vahou hlasu; každý z modelů v řadě se zaměřuje jen na ty případy, které předcházející modely nedokázaly správně klasifikovat. Při kombinování modelů se obvykle zvyšuje správnost klasifikace. V našem případě byla pomocí základního klasifikátoru založeného na pravidlech dosažena průměrná správnost u 30 realizovaných testů 70,20 %, při použití metody boosting to bylo 69,71 %. Nejlepší výsledek byl dosažen pomocí klasifikátorů vytvořených na základě algoritmu C5.0 s využitím metody boosting. Byla získána 76,65% celková přesnost. Nejhorší výsledky podle tab. 5 vykazuje algoritmus CHAID ($A_c = 61,71$ %).

Tab. 3: Průměrné hodnoty 30 testů – testovací data

Metoda	Celková chyba E_c [v %]	Celková přesnost A_c [v %]
C5.0	29,80	70,20
C5.0 boosting	30,29	69,71
C&RT	30,62	69,38
CHAID	31,17	68,83

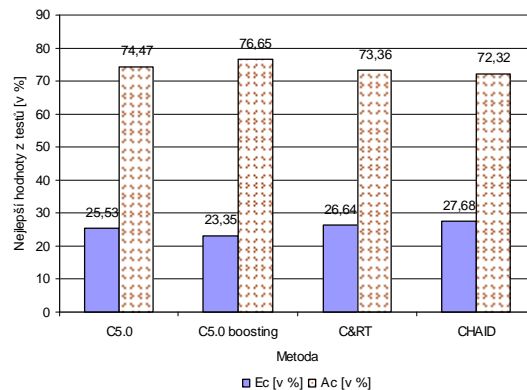
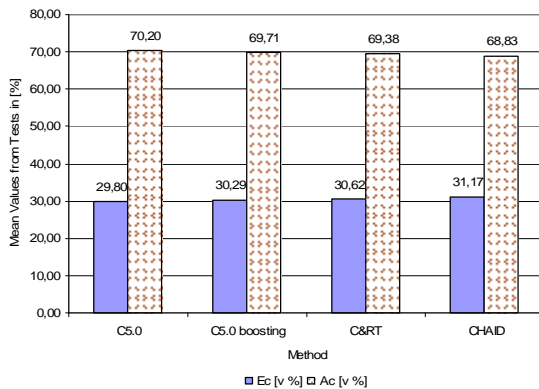
Tab. 4: Nejlepší hodnoty 30 testů – testovací data

Metoda	Celková chyba E_c [v %]	Celková přesnost A_c [v %]
C5.0	25,53	74,47
C5.0 boosting	23,35	76,65
C&RT	26,64	73,36
CHAID	27,68	72,32

Tab. 5: Nejhorší hodnoty 30 testů – testovací data

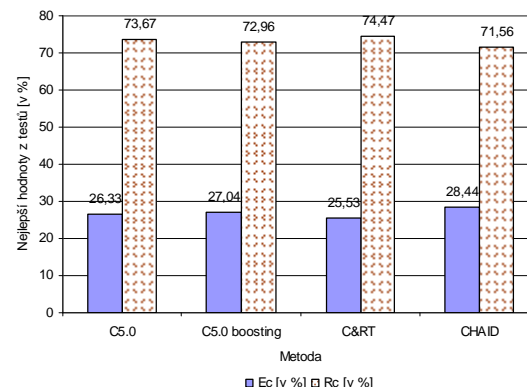
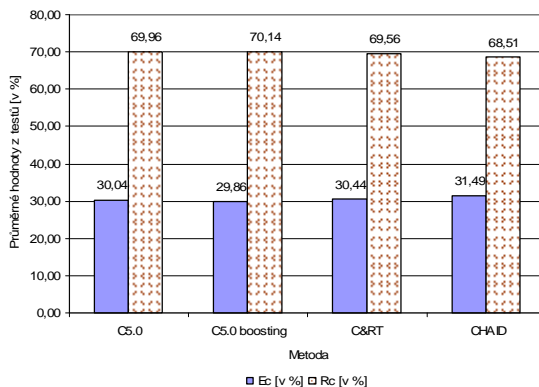
Metoda	Celková chyba E_c [v %]	Celková přesnost A_c [v %]
C5.0	35,24	64,76
C5.0 boosting	35,51	64,49
C&RT	36,24	63,76
CHAID	38,29	61,71

Grafické vyjádření výsledků jednotlivých klasifikátorů realizovaných na testovacích datech ukazují obr. 4 a obr. 5. Výsledky vytvořené na trénovacích datech, tzv. resubstituční chybu R_c , uvádí obr. 6 a obr. 7.



Obr.4: Průměrné hodnoty z 30 testů – testovací data

Obr.5: Nejlepší hodnoty z 30 testů – testovací data



Obr.6: Průměrné hodnoty z 30 testů – trénovací data

Obr.7: Nejlepší hodnoty z 30 testů – trénovací data

Na obr. 6 jsou zachyceny průměrné hodnoty R_c na trénovacích datech, obr. 7 znázorňuje nejlepší dosažené hodnoty klasifikátorů na základě uvedených algoritmů. Lze konstatovat, že algoritmy dosahují podobných výsledků.

3. Závěr

Spokojenost člověka je jednou ze základních podmínek určujících kvalitu jeho života. Je to však veličina značně subjektivní a mění 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 nejen s životním prostředím, ve kterém žijí, 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, a proto je třeba spokojenost občanů hodnotit a měřit.

Případová studie, která se zabývala modelováním spokojenosti občanů s kvalitou životního prostředí, byla pojata jako úloha klasifikační, v které bylo cílem zařazení občanů do tříd podle určení jejich spokojenosti. Byly zde navrženy modely klasifikátorů vytvořené na základě algoritmů C5.0, CHAID a C&RT. Nejlepších výsledků bylo dosaženo pomocí algoritmu C5.0. Avšak vzhledem k výsledkům ostatních modelů lze říci, že bylo dosaženo srovnatelných výsledků. Ke zvýšení přesnosti modelů se jeví vhodné využití dalších vstupních atributů (dotazníkových otázek) charakterizujících sledovanou problematiku.

Poděkování

Tento příspěvek vznikl za podpory projektu GAČR č. 402/08/0849 „Model řízení povzbudivého růstu regionu“.

Použitá literatura:

- [1] ADAMČÍK, S. *Regionální politika a management regionů, obcí a měst*. 1. vyd. Ostrava: VŠB - Technická univerzita Ostrava., 2000. ISBN: 80-7078-837-0.
- [2] BERKA, P. *Dobývání znalostí z databází*. 1. vyd. Praha: Academia, 2003. ISBN 80-200-1062-9.
- [3] BRYCHTOVÁ, Š., DUPLINSKÝ, J. Image obce - pohled občana. In *Aktuální otázky rozvoje regionů*. 1. vyd. Pardubice: Univerzita Pardubice, 2005. s. 57 - 64. ISBN 80-7194-777-6.
- [4] GUIDICI, P. *Applied Data Mining: Statistical Methods for Business and Industry*. West Sussex: Wiley, 2003. 376 s. ISBN 0-47084679-8.
- [5] HAN, J., KAMBER, M. *Data Mining: Concepts and Techniques*. Morgan Kaufmann Press, 2001. 770 s. ISBN 1-55860-901-6.
- [6] CHAPMAN, P. et al. *CRISP-DM 1.0: Step-by-step data mining guide* [online]. 15.8.2000 [cit. 2008-04-09]. Dostupné na WWW: <<http://www.crisp-dm.org/CRISPWP-0800.pdf>>.
- [7] JALOVECKÁ, M. 5. konference kvality ve veřejné správě – Paříž 2008. In *Veřejná správa: týdeník vlády České republiky*. 2008, roč. 11, č. 5, s. VII-VIII. ISSN: 1213-6581.
- [8] KRATOCHVÍL, I. *O řízení vážně i s úsměvem*. Kladno: Ing. Vladimír Macek, [2000]. ISBN 80-86091-32-5.
- [9] MAIMON, O., ROKACH, L. *Decomposition Methodology for Knowledge Discovery and Data Mining*. World Scientific Publishing, 2005. 323 s. ISBN 978-9812560797.
- [10] MV ČR. *Kvalita ve veřejné správě* [online]. c2005 [cit. 2008-04-09]. Dostupné na WWW: <<http://www.mvcr.cz/odbor/moderniz/koncepce/kvalita.html>>.
- [11] NOVÁK, J. *Evropské indikátory udržitelného rozvoje v praxi měst České republiky* [online]. 25.09.2006, poslední revize 1.2.2007 [cit. 2007-09-02]. Dostupné na WWW: <http://www.timur.cz/index2.php?option=com_docman&gid=13&task=doc_view&Itemid=38>.
- [12] PÍRKO, J. *Modelování pravidlových systémů v oblasti veřejné správy*. Fakulta ekonomicko-správní, Univerzita Pardubice: Pardubice, 2008, 80 s.

- [13] PS. 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 na WWW: <http://www.reccr.cz/download/indikator/technicka_zprava_cz.rtf>.
- [14] PŮČEK, M. et al. *Měření spokojenosti v organizacích veřejné správy -: 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 na WWW: <http://www.mvcr.cz/odbor/moderniz/spokojenost_final.pdf>.
- [15] PYLE, D. *Business Modeling and Data Mining*. Morgan Kaufmann Publishers, 2003. 693 s. ISBN 1-55860-653-X.
- [16] RYCHLÝ, M. *Klasifikace a predikce* [online]. 11.1.2006 [cit. 2007-11-22]. Dostupné na WWW: <<http://www.fit.vutbr.cz/~rychly/docs/classification-and-prediction/classification-and-prediction.pdf>>.
- [17] SHINNERS, S. M. *Modern Control System Theory and Design*. Second Edition, John Wiley and Sons, Inc. : New York, 1998. ISBN 0-471-24906-8.
- [18] SPSS. *SPSS Classification Trees* [online]. 5. 2. 2008 [cit. 2007-11-22]. Dostupné na WWW: <http://www.spss.cz/sw_mcla.htm>.
- [19] SUR. *Strategie udržitelného rozvoje České republiky* [online]. 21.12.2004 [cit. 2007-12-06]. Dostupné na WWW: <[http://www.env.cz/AIS/web-pub.nsf/\\$pid/MZPISF7Z6L7V](http://www.env.cz/AIS/web-pub.nsf/$pid/MZPISF7Z6L7V)>
- [20] ŠPIRKO, Š., KŘUPKA, J. *Základy technickej kybernetiky*. Akadémia ozbrojených síl M. R. Štefánka v Liptovskom Mikuláši : Liptovský Mikuláš, 2008, 283 s. ISBN 978-80-8040-357-7.
- [21] TER-MANUELIANC, A. *Moderní technologie řízení I*. Institut řízení : Praha, 1990. ISBN 80-7014-022-4.
- [22] 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.
- [23] VORLÍČEK, Z. Strategie Národní politiky kvality v České republice na období let 2008 až 2013 pro vyšší kvalitu života občanů České republiky. *In Veřejná správa: týdeník vlády České republiky*. 2008, roč. 11, č. 5, s. III-V. ISSN: 1213-6581.
- [24] WITTEN, I.H., FRANK, E. *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufman, 2005. 526 s. ISBN 0-12-088407-0
- [25] 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

Kontaktní adresy:

Ing. Miloslava Kašparová, Ph.D.
 Ústav systémového inženýrství a informatiky
 Fakulta ekonomicko-správní
 Univerzita Pardubice
 Studentská 95
 532 10 Pardubice
 e-mail: Miloslava.Kasparova@upce.cz
 tel.: 466036245

doc. Ing. Jiří Křupka, CSc.
Ústav systémového inženýrství a informatiky
Fakulta ekonomicko-správní
Univerzita Pardubice
Studentská 95
532 10 Pardubice
e-mail: Jiri.Krupka@upce.cz
tel.: 466036515

Ing. Jaromír Pírko
Staré Náměstí 823
53821 Slatiňany
e-mail: Pirko.Jaromir@seznam.cz