

Česká zemědělská univerzita v Praze

Provozně ekonomická fakulta

Katedra statistiky



Diplomová práce

Klasifikace dat z průzkumu "Veřejnost jako aktér modernizace" s využitím diskriminační analýzy a metody CART

Bc. Jaroslav Dlabač

© 2017 ČZU v Praze

ZADÁNÍ DIPLOMOVÉ PRÁCE

Jaroslav Dlabač

Projektové řízení

Název práce

Klasifikace dat z průzkumu "Veřejnost jako aktér modernizace" s využitím diskriminační analýzy a metody CART

Název anglicky

Data classification from the "Public as an actor of modernization" survey using discriminatory analysis and the CART method

Cíle práce

Hlavním cílem diplomové práce bude klasifikace datového souboru "Veřejnost jako aktér modernizace" získaného z Českého sociálně vědního datového archivu Sociologického ústavu AV ČR. Student bude aplikovat diskriminační analýzu a metodu klasifikačních a regresních stromů a následně určí, která z těchto dvou metod, z nichž každá představuje odlišný přístup ke klasifikaci dat, je pro danou analýzu vhodnější. Student se bude v rámci svého diplomového úkolu zabývat analýzou postojů a perspektivního myšlení českých elit a veřejnosti jako aktérů modernizace v závislosti na jejich hodnotovém zázemí a vývoji vzájemné interakce. Bude hledat odpovědi na otázky: Jaké vize moderní společnosti a preferované strategie modernizace ČR jsou charakteristické pro různé kategorie a skupiny veřejnosti. Jak ovlivňuje hodnotové zázemí aktérů variabilitu preferovaných vizí a strategií? Do jaké míry korespondují hodnotové orientace aktérů s jejich představami o strategiích modernizace?

Podle hlavního tvrzení teorie modernizace je ekonomický rozvoj spojen se souvislými a do jisté míry předvídatelnými změnami v kultuře a společenském a politickém životě. Výsledky analýz celosvětových průzkumů nasvědčují skutečnosti, že rozvoj hospodářství vede společnosti zhruba předvídatelným směrem: Industrializace vede k oborové specializaci, rostoucí úrovni vzdělanosti, rostoucím příjmům a nakonec přináší nečekané změny – v genderových rolích, v postojích k autoritě, pokles porodnosti ...

Metodika

Zdrojem dat budou kvantitativní sociologické výzkumy. Datové soubory student získá z Českého sociálně vědního datového archivu Sociologického ústavu AV ČR. Pro jejich analýzu použije vícerozměrné statistické metody umožňující odhalení závislostí a následnou klasifikaci proměnných. Předpokládá se aplikace diskriminační analýzy a metody CART (metody klasifikačních a regresních stromů), která patří k novějším metodám tzv. data miningu. Výsledky obou použitých metod budou porovnány a interpretovány.

Doporučený rozsah práce

60 – 80 stran

Klíčová slova

Klasifikační úloha, vícerozměrné statistické metody, diskriminační analýza, analýza, statistická analýza, modernizace, CART, data mining, rozhodovací stromy.

Doporučené zdroje informací

BERKA, P. Dobývání znalostí z databází. Praha: Academia, 2003. ISBN 80-200-1062-9.

Berry, M. J. A. – Linoff, G. Data mining techniques for marketing, sales, and customer relationship management. Indianapolis: Wiley, 2004. ISBN 0-471-47064-3.

HEBÁK, P., JAROŠOVÁ, E., PECÁKOVÁ, I., PLAŠIL, M., ŘEZANKOVÁ, H., VILIKUS, O., VLACH, P. Statistické myšlení a nástroje analýzy dat. 1. vyd. Praha: Informatorium, 2013. 877 s. ISBN 978-80-7333-105-4.

KELLER, J. Teorie modernizace. Praha: Slon, 2007. ISBN: 978-80-86429-66-3.

PECÁKOVÁ, I. Statistika v terénních průzkumech. 2. vyd. Praha: PROFESSIONAL PUBLISHING, 2011. 236 s. ISBN 978-80-7431-039-3.

SPSS INC. SPSS statistical algorithms. Chicago: SPSS, 1991. ISBN 0-918469-89-9.

Předběžný termín obhajoby

2017/18 ZS – PEF (únor 2018)

Vedoucí práce

Ing. Radka Procházková, Ph.D.

Garantující pracoviště

Katedra statistiky

Elektronicky schváleno dne 25. 11. 2016

prof. Ing. Libuše Svatošová, CSc.

Vedoucí katedry

Elektronicky schváleno dne 25. 11. 2016

Ing. Martin Pelikán, Ph.D.

Děkan

V Praze dne 29. 11. 2017

Čestné prohlášení

Prohlašuji, že svou diplomovou práci " Klasifikace dat z průzkumu "Veřejnost jako aktér modernizace" s využitím diskriminační analýzy a metody CART" jsem vypracoval samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou citovány v práci a uvedeny v seznamu použitých zdrojů na konci práce. Jako autor uvedené diplomové práce dále prohlašuji, že jsem v souvislosti s jejím vytvořením neporušil autorská práva třetích osob.

V Praze dne 30.11.2017

Poděkování

Rád bych touto cestou poděkoval Ing. Radce Procházkové, Ph.D. za její připomínky, konzultace a cenné rady, které mi při psaní této diplomové práce poskytla. Zároveň bych jí chtěl velmi poděkovat za její milý a vstřícný přístup.

Klasifikace dat z průzkumu "Veřejnost jako aktér modernizace" s využitím diskriminační analýzy a metody CART

Abstrakt

Tato diplomová práce se zabývá klasifikací dat z datového souboru „Veřejnost jako aktér modernizace“ získaného z Českého sociálněvědního datového archivu. Pro klasifikaci budou použity dvě metody, z nichž každá z nich zastává rozdílný přístup v klasifikaci dat. První metodou je diskriminační analýza, která představuje základní statistickou metodu pro vícerozměrnou statistiku. Druhá metoda je metoda klasifikačních a regresních stromů, známá též jako metoda CART, která představuje nové postupy v oblasti data miningu. Dílčí analýza je zaměřena na hledání závislost v datovém souboru a snaží se odpovědět na otázku, zda je možné osoby, které jsou pro modernizaci, klasifikovat podle určitých znaků, jako například pohlaví, vzdělání nebo rodinný stav a zda je možné potvrdit některé poznatky řečené v teoretické části.

Klíčová slova: klasifikace, diskriminační analýza, metoda CART, rozhodovací stromy, data mining, modernizace

Data classification from “The Public as an actor of modernization” survey using discriminatory analysis and the CART method

Abstract

This diploma thesis deals with the classification of data from the data set "The public as an actor of modernization" obtained from the Czech Socio-scientific data archive. For classification are used two methods, each with different approach in data classification. The first method is discrimination analysis, which is the basic statistical method for multidimensional statistics. The second method is method of classification and regression trees, also known as the CART method, which is using new practices in data mining. The partial analysis is aimed at finding dependence on a data file and attempts to answer the question of whether it is possible to classify people who agree modernization by certain basic features such as gender, education, or marital status and whether it is possible to confirm some of the knowledge mentioned in the theoretical part

Keywords: classification, discrimination analysis, CART method, decision trees, data mining, modernization

OBSAH

1	ÚVOD	4
2	CÍL PRÁCE A METODIKA	6
2.1	CÍL PRÁCE	6
2.2	METODIKA PRÁCE	7
2.2.1	<i>Popis klasifikační úlohy</i>	7
2.2.2	<i>Technologický pohled na dobývání znalostí z databází</i>	7
2.2.3	<i>Manažerský pohled na dobývání znalostí z databází</i>	8
2.2.4	<i>Cíl klasifikační úlohy</i>	9
2.2.5	<i>Algoritmus obecné klasifikační úlohy</i>	9
2.3	DISKRIMINAČNÍ ANALÝZA.....	12
2.3.1	<i>Podstata a cíl diskriminační analýzy</i>	12
2.3.2	<i>Předpoklady diskriminační analýzy</i>	13
2.3.3	<i>Kanonická diskriminační analýza</i>	14
2.3.4	<i>Diskriminace mezi dvěma skupinami</i>	14
2.3.5	<i>Klasifikace v případě tří a více skupin</i>	15
2.3.6	<i>Vyhodnocení diskriminační analýzy</i>	15
2.4	METODA KLASIFIKAČNÍCH A REGRESNÍCH STROMŮ	17
2.4.1	<i>Základní algoritmus rozhodovacího stromu</i>	17
2.4.2	<i>Prořezávání stromů</i>	18
2.4.3	<i>Regresní stromy</i>	19
2.4.4	<i>Metoda CART</i>	21
2.4.5	<i>Použití rozhodovacích stromů</i>	22
3	TEORETICKÁ VÝCHODISKA	24
3.1	PRVNÍ VLNA MODERNIZACE	26
3.2	DRUHÁ VLNA MODERNIZACE	28
3.3	GLOBALIZACE	31
3.3.1	<i>První fáze globalizace</i>	31
3.3.2	<i>Druhá fáze globalizace</i>	33
4	VLASTNÍ PRÁCE	36
4.1	POPIS DATOVÉHO SOUBORU	36
4.2	ANALÝZA ZÁVISLOSTI PŘEMĚNY SPOLEČNOSTI NA TEMPU MODERNIZACE	36
4.3	ANALÝZA GLOBALIZACE V RÁMCI MODERNIZACE.....	41
4.4	ANALÝZA POSTOJŮ K MODERNIZACI	45
4.5	ANALÝZA POSTOJŮ K EKONOMICKÉMU RŮSTU	48
4.6	ANALÝZA ZÁVISLOSTI PŘEMĚNY SPOLEČNOSTI NA TEMPU MODERNIZACE – METODA CART	51

4.7	ANALÝZA GLOBALIZACE V RÁMCI MODERNIZACE – METODA CART	54
4.8	ANALÝZA POSTOJŮ K MODERNIZACI – METODA CART	57
4.9	ANALÝZA POSTOJŮ K EKONOMICKÉMU RŮSTU – METODA CART	60
4.10	POROVNÁNÍ VÝSLEDKŮ OBOU METOD	64
5	ZÁVĚR.....	66
6	SEZNAM ZDROJŮ	68
7	PŘÍLOHY	70

Seznam obrázků

Algoritmus obecné klasifikační úlohy	9
Příklad regresního stromu	20
Příklad rozhodovacího stromu metody CART	21
Příklad vyjadřovací síly rozhodovacích stromů.....	23
Výstup z programu IBM SPSS – První dvě úrovně v CART – analýza č.1	52
Výstup z programu IBM SPSS – Pravá větev – analýza č.1.....	53
Výstup z programu IBM SPSS – První dvě úrovně CART – analýza č.2	55
Výstup z programu IBM SPSS – Další úrovně v metodě CART – analýza č.2	56
Výstup z programu IBM SPSS – První dvě úrovně metody CART – analýza č.3	58
Výstup z programu IBM SPSS – Další úrovně metody CART – analýza č.3	59
Výstup z programu IBM SPSS – První dvě úrovně metody CART – analýza č.4	61
Výstup z programu IBM SPSS – Levá větev metody CART – analýza č.4	62
Výstup z programu IBM SPSS – Pravá větev metody CART – analýza č.4.....	63
Výstup z programu IBM SPSS Levá a prostřední větev – analýza č. 1	74

Seznam tabulek

Výstup z programu IBM SPSS – Analýza proměnných v analýze č.1	37
Výstup z programu IBM SPSS – Wilksovo lambda rozdělení v analýze č.1	37
Výstup z programu IBM SPSS – Wilksovo lambda rozdělení v analýze č.1	38
Výstup z programu IBM SPSS – Kanonická korelace v analýze č.1	39
Výstup z programu IBM SPSS – Kanonická struktura v analýze č.1	39
Výstup z programu IBM SPSS – Klasifikace do jednotlivých skupin v analýze č.1	40

Výstup z programu IBM SPSS – Analýza proměnných v analýze č.2.....	41
Výstup z programu IBM SPSS – Testování diskriminační schopnosti v analýze č.2	42
Výstup z programu IBM SPSS – Wilksovo lambda rozdělení v analýze č.2.....	42
Výstup z programu IBM SPSS – Kanonická korelace v analýze č.2	43
Výstup z programu IBM SPSS – Kanonická struktura v analýze č.2.....	43
Výstup z programu IBM SPSS – Klasifikace do jednotlivých skupin v analýze č.2	44
Výstup z programu IBM SPSS – Analýza proměnných v analýze č.3.....	45
Výstup z programu IBM SPSS – Testování diskriminační schopnosti v analýze č.3	46
Výstup z programu IBM SPSS – Wilksovo lambda rozdělení v analýze č.3	46
Výstup z programu IBM SPSS – Kanonická korelace v analýze č.3	47
Výstup z programu IBM SPSS – Klasifikace do jednotlivých skupin v analýze č.3	47
Výstup z programu IBM SPSS – Analýza proměnných v analýze č.4.....	48
Výstup z programu IBM SPSS – Testování diskriminační schopnosti v analýze č.4	49
Výstup z programu IBM SPSS – Wilksovo lambda rozdělení v analýze č.4.....	49
Výstup z programu IBM SPSS – Kanonická korelace v analýze č.4	50
Výstup z programu IBM SPSS – Klasifikace do jednotlivých skupin v analýze č.4	50
Výstup z programu IBM SPSS – Tabulka Risk – analýza č.1	54
Výstup z programu IBM SPSS – Klasifikace podle CART – analýza č.1	54
Výstup z programu IBM SPSS – Tabulka Risk – analýza č.2.....	57
Výstup z programu IBM SPSS – Klasifikace podle CART – analýza č.2	57
Výstup z programu IBM SPSS Tabulka Risk – analýza č.3.....	59
Výstup z programu IBM SPSS – Klasifikace podle CART – analýza č.3	60
Výstup z programu IBM SPSS – Tabulka Risk – analýza č.4.....	63
Výstup z programu IBM SPSS – Tabulka Risk – analýza č.4.....	64
Výstup z programu IBM SPSS – Rozložení respondentů v analýze č.1	70
Výstup z programu IBM SPSS – Rozložení respondentů v analýze č.2	71
Výstup z programu IBM SPSS – Rozložení respondentů v analýze č.3	72
Výstup z programu IBM SPSS Rozložení respondentů v analýze č.4	73

1 ÚVOD

Diskriminační analýza a metoda klasifikačních a regresních stromů, zkráceně metoda CART, patří k metodám, které jsou využívány k diskriminaci a následně i klasifikaci dat do určitých tříd. Zatímco diskriminační analýza patří k nejstarším metodám vícerozměrné statistiky, metoda CART patří k novým data miningovým metodám, která využívá tzv. učení s učitelem a umožňuje i relativním začátečníkům získat přehledný výstup z analýzy.

Sociologové, podnikatelé, ekonomové, politici a běžní občané všeobecně vnímají slovo „modernizace“ jako určitý proces, pomocí něhož se uvádí v podniku, ve strojírenství, IT či jiném odvětví daný proces, technologie nebo strategie na současný stav či trend. Co ovšem přímo souvisí se slovem modernizace, jsou modernizační teorie. Tyto teorie popisují progresivní proces, pomocí něhož je společnost převáděna z pre-moderního na tradiční a poté na moderní typ společnosti. Podle názorů odborníků by měl být modernizován také celý sociální stát. Procesem modernizace by měl také projít životní styl, hodnoty a postoje obyvatel.

V metodické části práce je nejprve popsána klasifikační úloha, její cíl a podstata. Dále zde je popsána diskriminační analýza, její princip, cíl, předpoklady a její různé typy. V další části je nejprve vysvětlen základní algoritmus rozhodovacího stromu a jeho následné prořezávání. Následně je věnována samostatná podkapitola metodě CART a celá metodická část je zakončena využitím klasifikačních stromů v praxi.

V teoretických východiscích práce je nejprve seznámení s pojmem modernizace a popsání první a druhé vlny modernizace, tak jak o nich mluví přední světoví autoři, zaměřující se na tuto problematiku. Na pojem modernizace je úzce navázán i pojem globalizace, jehož spojitost s modernizací je vysvětlena v dalších kapitolách stejně tak, jako její první a druhá vlna.

Praktická část práce je pak věnována analýzám, za použití výše zmíněných metod, na datovém souboru vztahujícím se právě k modernizace. Zde jsou nejprve provedeny 4 analýzy za pomoci diskriminační analýzy a jejich výsledky detailně popsány. Analýzy se týkají závislosti přeměny společnosti na tempu modernizace, globalizace v rámci modernizace, postojů k modernizaci a postojů k ekonomickému růstu. Následně jsou naprosto stejné analýzy provedeny pomocí metody CART a jejich výsledky jsou opět do

hloubky popsány. Poslední kapitola se pak věnuje porovnání výsledků těchto dvou metod a zhodnocení, která z vybraných metod přinesla v rámci klasifikace hodnotnější výsledek.

Veškeré analýzy byly realizovány pomocí programu IBM SPSS Statistics 20.

2 CÍL PRÁCE A METODIKA

2.1 Cíl práce

Cílem diplomové práce je klasifikace datového souboru „Veřejnost jako aktér modernizace“, který pro účely práce poskytl Český sociálněvědní datový archiv. Ke klasifikaci budou využity dvě rozdílné metody, a to sice diskriminační analýza a metoda klasifikačních a regresních stromů, zkráceně metoda CART. Každá z těchto dvou metod zastává rozdílný přístup. Zatímco diskriminační analýza je základní metody pro zjišťování závislosti a následné klasifikace proměnných u vícerozměrné statistiky, metoda CART zastává tzv. učení s učitelem. Následně pak bude určeno, která ze dvou vybraných metod je pro daný případ vhodnější a dává hodnotnější výsledky.

Dílním cílem bude hledání závislosti mezi proměnným v datovém souboru týkajícím se modernizace a názorem veřejnosti na ni. Práce bude hledat odpovědi na otázky, zdali je možné říci, kteří obyvatelé ČR jsou pro modernizaci a kteří ne. Dále pak jestli je možné najít závislost mezi modernizací a globalizací, která je jevem úzce souvisejícím s modernizací. Dále také jestli je možné vypořádat závislost mezi názory na společenskou změnu v naší společnosti a modernizací jako takovou. Podle hlavního tvrzení teorie modernizace je ekonomický rozvoj spojen se souvislými a do jisté míry předvídatelnými změnami v kultuře a společenském a politickém životě. Výsledky analýz celosvětových průzkumů nasvědčují skutečnosti, že rozvoj hospodářství vede společnosti zhruba předvídatelným směrem: Industrializace vede k oborové specializaci, rostoucí úrovni vzdělanosti, rostoucím příjmům, a nakonec přináší nečekané změny například v genderových rolích, v postojích k autoritě, pokles porodnosti.

2.2 Metodika práce

2.2.1 Popis klasifikační úlohy

Klasifikační úloha je spolu s predikcí, deskripcí a hledáním „nugettů“¹ jedna z typických úloh dobývání znalostí z databází, též KDD (Knowledge discovery in databases). Jejím cílem je nalézt nějakou určitou znalost pro třídění dat, která by se dala použít jako podklad do budoucna. Je zde dána přednost pokrytí na úkor jednoduchosti. Na samotné dobývání znalostí z databází je možné se dívat ze dvou pohledů – z technologického a manažerského.²

2.2.2 Technologický pohled na dobývání znalostí z databází

Technologický pohled je možno rozdělit do pěti částí.³

- 1) Selekcce – na začátku analýzy je dán výchozí soubor o velkém počtu proměnných, ze kterého je třeba vyselektovat vybraná data. Výsledkem je tabulka, jež bude obsahovat relevantní údaje pro daný výzkum.
- 2) Předzpracování – zde se již pracuje s menším datovým souborem. Tato část je časově nejnáročnější, protože je třeba odstranit odlehlé hodnoty a celkově vyčistit datový soubor.
- 3) Transformace – v tomto bodě dochází k nahrazení, též diskretizaci spojitych veličin, které byly odstraněny v přechozím bodě.
- 4) Data mining – samotný popis data miningu (dolování) by vydal na samostatnou kapitolu, ale ve zkratce se jedná o získávání skryté a pro analytika potencionálně užitečné informace z datového souboru. Zahrnuje to určitý soubor vybraných metod pro hledání pro nás zajímavých vztahů v datech. Obvykle se zde kombinuje několik metod na základě dílčího výsledku. Data mining bývá také někdy chybně zaměňován se samotným dobývání znalostí z databází, ale jedná se pouze o jeho součást.
- 5) Interpretace – jedná se o interpretaci výsledků zjištěných daným výzkumem uživateli. Ne vždy je však možné výsledky hned interpretovat a často je potřeba výsledky upravit do určité podoby, které bude koncový uživatel rozumět.⁴

¹ Nugetty jsou nové, zajímavé a překvapivé znalost, které nemusí plně pokrývat daný koncept.

² BERKA, P. *Dobývání znalostí z databází*. 2003 s 15-19

³ MRÁZOVÁ, Iveta. Dobývání znalostí. [online] a PLCHÚT, Martin. Dobývání znalostí z databází: Úvod a oblasti aplikací. [online]

2.2.3 Manažerský pohled na dobývání znalostí z databází

V manažerském pohledu je hlavním cílem získat co nejvíce relevantních informací pro řešení daného problému.

Samotný proces má na rozdíl od technologického pohledu 7 částí.⁵

- 1) Řešitelský tým – v samém počátku se vytvoří tým, který bude daný problém řešit. Členové týmu musí být experti na danou tematiku, experti na práci s daty nebo na metody dobývání znalostí s databází.
- 2) Specifikace problému – je přesně stanoven problém, který chce řešitelský tým vyřešit.
- 3) Získání dat – proces, který může být značně zdlouhavý, protože sehnat vhodná data je ne vždy snadné. Často se tak pracuje s daty, která už jsou poněkud zastaralá. Za určitých okolností to může vést i k přeformulování problému nebo jeho drobné změně. V mnohých případech se tedy volí metoda, kde jsou použita externí data popisující prostřední, v němž se analyzované děje odehrávají.
- 4) Výběr metod – zde se vybere vhodná metoda na analýzu dat. Jedná se o různé klasifikační metody. V případě této diplomové práce se jedná o diskriminační analýzu a metodu klasifikační a rozhodovacích stromů.
- 5) Předzpracování dat – v podstatě se jedná o stejný krok jako v technologickém postupu čili je třeba připravit si data do vhodné formy pro vybrané metody. Často se může jednat o náročné početní operace.
- 6) Data mining – opět stejná část jako v předchozím postupu. Aplikace zvolených analytických metod. Obvykle jsou metody aplikovány víckrát. Hodnoty vstupních parametrů jednotlivých běhů závisí na výsledcích předchozích běhů.
- 7) Interpretace – zde je potřeba upravit výsledky do formy, která bude srozumitelná pro uživatele. Například ve formě analytické zprávy.⁶

⁴ MRÁZOVÁ, Iveta. Dobývání znalostí. [online] a PLCHÚT, Martin. Dobývání znalostí z databází: Úvod a oblasti aplikací. [online]

⁵ MRÁZOVÁ, Iveta. Dobývání znalostí. [online]

⁶ MRÁZOVÁ, Iveta. Dobývání znalostí. [online]

2.2.4 Cíl klasifikační úlohy

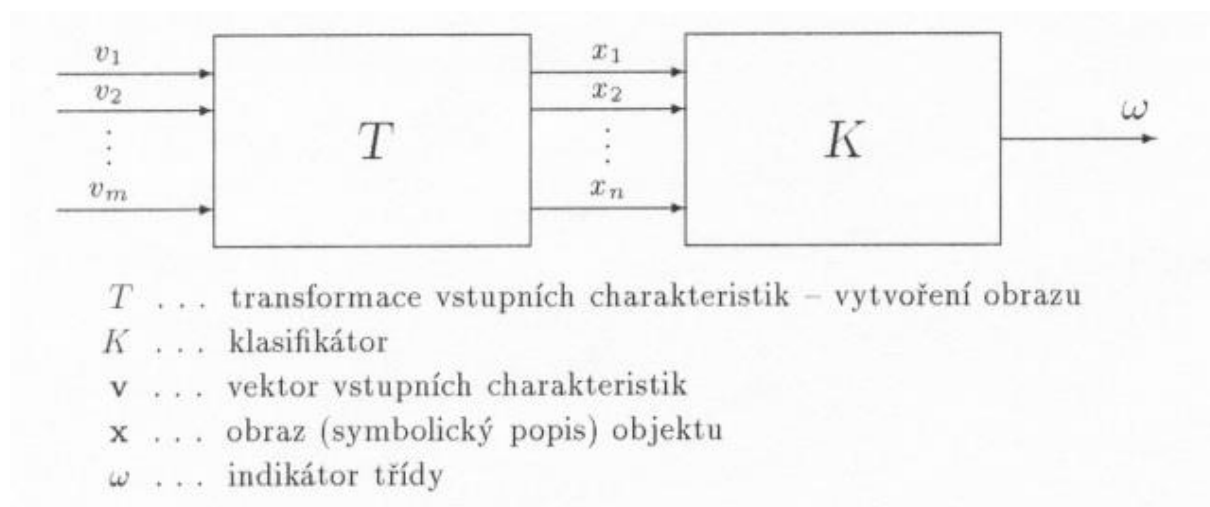
Cílem klasifikační úlohy je každý jednotlivý případ v souboru zařadit do konkrétní odhadnuté třídy. Obecně je klasifikační úloha založena na výpočtu podmíněných pravděpodobností jednotlivých zařazení. Metody, při kterých jsou rozpoznány jednotlivé případy, je možné rozdělit do tří skupin, a to sice na klasické statistické metody, metody založené na strojovém učení a metody založené na umělých neuronových sítích tzv. neuronové klasifikátory.⁷

2.2.5 Algoritmus obecné klasifikační úlohy

Algoritmus obecné klasifikační úlohy je v podstatě totožný s technologickým pohledem, který je popsán ve výše uvedené kapitole. Opět je zde hledán určitý klasifikátor, pomocí něhož je možné rozřazovat proměnné do jednotlivých kategorií.

Graficky by se dal algoritmus obecné klasifikační úlohy zobrazit takto – viz obrázek č. 1.⁸

Obrázek 1: Algoritmus obecné klasifikační úlohy



Zdroj: BERKA, P. Dobývání znalostí z databází

Klasifikátor neboli rozpoznávací znaky je možné hledat třemi způsoby.⁹

- 1) Pomocí klasických statistických metod – jedná se o klasické statistické metody pro vícerozměrnou statistiku, jako například kontingenční tabulka, regresní

⁷ PLCHÚT, Martin. Dobývání znalostí z databází: Úvod a oblasti aplikací. [online]

⁸ PLCHÚT, Martin. Dobývání znalostí z databází: Úvod a oblasti aplikací. [online]

⁹ MRÁZOVÁ, Iveta. Dobývání znalostí. [online] a PLCHÚT, Martin. Dobývání znalostí z databází: Úvod a oblasti aplikací. [online]

analýza, diskriminační analýza, které je věnována pozdější kapitola a shluková analýza.

- 2) Pomocí metod založených na strojovém učení – tyto metody se dělí na učení s učitelem (supervised learning) a učení bez učitele (unsupervised learning). Rozdíl mezi to dvěma metodami je, že v učení s učitelem je známa správná klasifikace, zatímco v učení bez učitele nikoli.
- 3) Pomocí neuronových klasifikátorů – tato metoda využívá znalosti mozkových nervových buněk (neuronů), kdy přijímá kladné a záporné podněty od jiných neuronů a ve chvíli, kdy počet těchto neuronů překročí určitý práh, se sám aktivuje.¹⁰

2.2.5.1 Vyhodnocení klasifikační úlohy

Vyhodnocení výsledků klasifikační úlohy patří k nejdůležitějším částem celého procesu. Výsledky šetření musí být užitečné, zajímavé, a hlavně srozumitelné pro konečné uživatele. K tomu, aby byly výsledky srozumitelné, jsou využívány grafické pomůcky, jako například křivky, grafy a další vizualizační metody. Při samotném vyhodnocování klasifikační úlohy se obvykle postupuje metodou učení s učitelem a využíváme metody testování modelů, které jsou založeny na testování nalezených znalostí na datech. Tato metoda umožňuje porovnat, jak dobře nalezené znalosti shodují s informací od učitele. Zvolené testovací metody jsou vybrány podle toho, jaká data jsou použita pro učení a jaká pro testování.¹¹

Dalším důležitým pojmem při vyhodnocování klasifikační úlohy je celková správnost. Profesor Berka (2003) k celkové správnosti píše: „*Celková správnost (overall accuracy) někdy též označovaná úspěšnost (successfulness) nebo komplementární celková chyba (overall error) je nejjednodušší charakteristika toho, jak jsou získané znalosti kvalitní či ne.*

Celková správnost se počítá jako

$$Acc = \frac{TP + TN}{TP + TN + FP + FN}, \quad (2.1)$$

¹⁰ MRÁZOVÁ, Iveta. Dobývání znalostí. [online] a PLCHÚT, Martin. Dobývání znalostí z databází: Úvod a oblasti aplikací. [online]

¹¹ BERKA, P. *Dobývání znalostí z databází. 2003 s 223-224*

kde *TP* (*true positive*) je počet příkladů, které systém správně zařadil do pozitivní třídy, *FP* (*false positive*) je pak počet příkladů, které systém naopak chybně zařadil do pozitivní třídy. *TN* (*true negative*) zase správně zařazené do negativní třídy a *FN* (*false negative*) špatně zařazené do negativní třídy. Například když je dán systém, který hodnotí bonitu klientů banky za účelem rozhodnutí o úvěru. Dopustí se chyby tehdy, když doporučí půjčku klientovi, který ji nesplatí, nebo situace, že zamítne půjčku klientovi, který by ji splatil.

Celková chyba se poté vypočítá následovně

$$Err = \frac{FP + FN}{TP + TN + FP + FN} \quad (2.2)$$

A v případě, že je potřeba zjistit počet správných nebo nesprávných, rozhodnutí systému je

$$Err = 1 - Acc. \text{ „}^{12} \quad (2.3)$$

2.2.5.2 Spolehlivost klasifikace

Jedná se o hodnocení klasifikace, které je založeno na metaučení. Metaučení (metalearning) je postup, při kterém se na výsledky klasifikátoru opětovně použije nějaký učící se algoritmus.¹³

Profesor Berka (2003) ve své práci uvádí: „Z původních trénovacích dat se vytvoří model, který je použit pro klasifikaci těchto trénovacích dat. Poté jsou vytvořena nová trénovací a testovací data tak, že k původním datům je přidán nový binární atribut vyjadřující, zda při klasifikaci došlo nebo nedošlo ke shodě se skutečností. Pro nová trénovací data je vytvořen model, který bude pro daný příklad předpovídat, zda dojde ke shodě. Model je pak testován na nových testovacích datech.“¹⁴

Zatímco v první metodě se hodnotila znalost pouze na základě jednoho testování, respektive jednoho čísla, v tomto případě se jedná o komplexnější popis chování modelu. Dalšími metodami, které se při hodnocení opírají o více testů, jsou například křivka učení, v ekonometrii pak metoda DEA a také křivka ROC, ta v této práci také je popsána.

¹² Citace BERKA, P. *Dobývání znalostí z databází. 2003 s 227*

¹³ BERKA, P. *Dobývání znalostí z databází. 2003 226-209*

¹⁴ Citace BERKA, P. *Dobývání znalostí z databází. 2003 s 229*

Je třeba zmínit, že všechny tyto metody používají i grafické znázornění výsledků testování.¹⁵

Se spolehlivostí klasifikace je úzce spojený i tzv. Křivka ROC¹⁶. Tato křivka pracuje se statistikami TP a FP a podobně jako celková správnost používá kritérium $TP[\%] = (1 - FP[\%])$. Používá se u modelů, kde je klasifikace doplňována pravděpodobností.

Samotná křivka se tvoří tak, že mění práh, při kterém je výsledek klasifikace interpretován jako pozitivní. Touto změnou prahu je možné simulovat chování modelu v případě změny poměru mezi počty příkladů obou třídy i změny cen za chybou klasifikaci, a na cenu chyb. Tato metoda obnáší i grafické znázornění, kdy na ose x máme procentuálně vyjádřenou FP a na ose y procentuální TP. Cílem tedy je, se v grafu pohybovat co nejbližší bodu (0,1), který je „vlevo nahoře“, a který nám znázorňuje bezchybnou klasifikaci.¹⁷

Pokud se nechtějí klasifikovat příklady do třídy, ale do numerických predikcí, jsou použity metody statistického kritéria správnosti, jako jsou střední kvadratická chyba, její odmocnina, střední absolutní chyba, relativní kvadratická chyba nebo korelační koeficient.¹⁸

2.3 Diskriminační analýza

Diskriminační analýza je jedna z metod vícenásobné statistické analýzy, jejímž cílem je identifikace proměnných (klasifikátorů) diskriminujících mezi předem danými skupinami objektů a následně klasifikovat dané objektů do skupin. Tato metoda pochází z 30. let 20. století a byla formulována R. A. Fischerem.¹⁹

2.3.1 Podstata a cíl diskriminační analýzy

Profesor Hebák (2013) píše: „*Předpokládejme, že vícerozměrná jednofaktorová analýza rozptylu vede k zamítnutí hypotézy o shodě vektorů středních hodnot $\mu_1 = \mu_2 = \dots = \mu_H$ v H skupinách, na něž se rozpadá stejně jako uvažovaný základní soubor i*

¹⁵ BERKA, P. *Dobývání znalostí z databází. 2003 s 233-234*

¹⁶ ROC znamená receiver operating characteristic. V překladu pracovní charakteristika přijímače.

¹⁷ BERKA, P. *Dobývání znalostí z databází. 2003 s 234*

¹⁸ BERKA, P. *Dobývání znalostí z databází. 2003 s 234*

¹⁹ PECÁKOVÁ, I. *Statistika v terénních průzkumech. 2011*

náhodný výběr (o rozsahu n) z něj pocházející. Takový výsledek jsme interpretovali jako závislost p sledovaných kvantitativních proměnných X_1, X_2, \dots, X_p na použitém třídícím kritériu U , a tedy nejčastěji na veličině nominální.“²⁰

A Iva Pecáková (2011) dodává: „Liší-li se skupiny jednotek prokazatelně v úrovni sledovaných kvantitativních veličin, lze si položit otázku, nakolik je uvažovanými veličinami ovlivněna příslušnost jednotky k určité skupině. Uvedenou souvislost lze rovněž chápat jako určité pravidlo, vedoucí k zařazení jednotek v souboru do skupin na základě zjištěných hodnot několika kvantitativních proměnných. Diskriminační analýza pak velmi často směřuje ke klasifikaci jednotek s neznámou skupinovou příslušností (jejichž zařazení se kupříkladu týká nějaký budoucí události). Konkrétní podoby klasifikační pravidlo nabývá na základě hodnot souboru kvantitativních proměnných a skupinové příslušnosti u těch jednotek, u nichž jsou všechny potřebné údaje k dispozici. Použije se pak třídění nezařazených jednotek.“²¹

Odvětví, kde se používá diskriminační analýza, jsou například medicína, biologie a v technické obory.²²

2.3.2 Předpoklady diskriminační analýzy

Klasifikace objektů při diskriminační analýze je založena na předpokladech o vlastnostech klasifikačních objektů. Je zde předpoklad vícerozměrného normálního rozdělení prediktorů v každé ze skupin. Může být tedy citlivá na přítomnost odlehlých hodnot.²³

Profesor Petr Berka (2003) k předpokladům píše: „Dále při diskriminační analýze předpokládáme, že ke každé třídě (hodnotě nominální veličiny) $c_t, t=1, \dots, T$, existuje (diskriminační) funkce f_t , taková, že

$$f_t(X) = \max_{k=1, \dots, T} f_k(X), \quad (2.4)$$

²⁰ Citace HEBÁK, P. -- JAROŠOVÁ, E. -- PECÁKOVÁ, I. -- PLAŠIL, M. -- ŘEZANKOVÁ, H. -- VILIKUS, O. -- VLACH, P. *Statistické myšlení a nástroje analýzy dat 2013 s 179*

²¹ Citace PECÁKOVÁ, I. *Statistika v terénních průzkumech. 2011 s 196*

²² HEBÁK, P. -- JAROŠOVÁ, E. -- PECÁKOVÁ, I. -- PLAŠIL, M. -- ŘEZANKOVÁ, H. -- VILIKUS, O. -- VLACH, P. *Statistické myšlení a nástroje analýzy dat 2013 s 179*

²³ Jarkovský, J. – Littnerová, S. *Vícerozměrné statistické metody [online]*

právě když příklad $X = [x_1, x_2, \dots, x_m]$ patří do třídy c_i . Za předpokladu normálního průběhu se hledání diskriminační funkce omezuje na odhad střední hodnoty μ_i (na základě výběrových parametrů) a kovariančních matic S_i (na základě výběrových rozptylů). Dále je vhodné, když mají prediktory v jednotlivých skupinách shodné kovarianční matice. Tento předpoklad lze ale odpustit.²⁴

2.3.3 Kanonická diskriminační analýza

Za cíl kanonické diskriminační analýzy je možné označit snahu najít lineární kombinaci p našich sledovaných proměnných $Y = v^T x$, kde $v^T = [v_1, v_2, \dots, v_p]$ je vektor parametrů

$x^T = [X_1, X_2, \dots, X_p]$, aby lépe, než jakákoliv lineární kombinace oddělovala K skupin tak, aby vnitroskupinová variabilita byla co nejmenší a meziskupinová variabilita co největší. Cílem tedy je nalézt maximum tzv. Fisherova diskriminačního kritéria

$$\lambda = \frac{Q_B(Y)}{Q_E(Y)} = \frac{v^T B v}{v^T E v}, \quad (2.5)$$

kde matice B vyjadřuje meziskupinovou a matice E vnitroskupinovou variabilitu původních proměnných. V diskriminačním kritériu potřeba maximalizovat vektor v_1 , který odpovídá největšímu charakteristickému číslu λ_1 matice BE^{-1} . Lineární kombinaci $v_1^T x = Y_1$ má název první diskriminant nebo také první kanonická proměnná. Jestliže jsou proměnné roztrženy pouze do dvou skupin, stačí pouze jeden diskriminant. V případě tří a více skupin je potřeba R diskriminantů.²⁵

2.3.4 Diskriminace mezi dvěma skupinami

Tento typ diskriminace nejlépe popisuje opět pan Petr Berka (2003) ve své knize:²⁶ „V případě lineární diskriminační analýzy mezi dvěma skupinami mají diskriminační funkce $f(t)$ podobu lineární kombinace.

$$f_t(\mathbf{X}) = q_{0t} + q_{1t}x_1 + q_{2t}x_2 + \dots + q_{mt}x_m \quad (2.6)$$

²⁴ Citace BERKA, P. *Dobývání znalostí z databází. 2003 s 53a 54*

²⁵ PECÁKOVÁ, I. *Statistika v terénních průzkumech. 2011 s 197*

²⁶ Citace BERKA, P. *Dobývání znalostí z databází. 2003 s 53*

V případě, že máme dvě třídy, hledáme funkci

$$f_i(\mathbf{X}) = f_1(\mathbf{X}) - f_2(\mathbf{X}) \quad (2.7)$$

Zvolené příklady poté hodnotíme podle znaménka této funkce. Můžeme si to představit jako přímku, která odděluje obě třídy od sebe.

Pokud chceme zaručit optimální klasifikaci ve smyslu minimální chyby, musíme použít jako diskriminační funkce podmíněné (aposteriori) pravděpodobnosti zařazení pozorování X do třídy c_i .

$$P(c_i | X) = \frac{P(X | c_i)P(c_i)}{\sum_k P(X | c_k)P(c_k)}. \quad (2.8)$$

Pro variantu se dvěma třídami bude vzorec vypadat následovně

$$f(X) = f_1(X) - f_2(X) = P(X | c_1)P(c_1) - P(X | c_2)P(c_2). \quad (2.9)$$

2.3.5 Klasifikace v případě tří a více skupin

Za předpokladu vícerozměrné normality a shody kovariančních matic je rozšířeno diskriminační kritérium pro H skupin, kde $H > 2$. Kritérium pro H -tou skupinu je dáno následovně

$$x^t \sum^{-1} \mu_h - \frac{1}{2} \mu_h^T \sum^{-1} \mu_h + \ln \pi_h, \quad h = 1, 2, \dots, H. \quad (2.10)$$

Jednotka neznámého původu s hodnotami pozorovaných proměnných x je potom zařazena do skupiny s nejvyšší hodnotou diskriminačního kritéria. Neznámé vektory průměrů μ_h a kovarianční matice společnou výběrovou kovarianční maticí S ; apriorní pravděpodobnosti π_h jsou odhadnuty výběrovými podíly jednotlivých skupin.²⁷

2.3.6 Vyhodnocení diskriminační analýzy

Profesor Hebák (2013) k vyhodnocení píše: „Při vyhodnocování zvoleného diskriminačního kritéria je pro nejpodstatnější pravděpodobnost mylné klasifikace

²⁷ HEBÁK, P. -- JAROŠOVÁ, E. -- PECÁKOVÁ, I. -- PLAŠIL, M. -- ŘEZANKOVÁ, H. -- VILIKUS, O. -- VLACH, P. *Statistické myšlení a nástroje analýzy dat 2013 s 204*

$$\sum_{h=1}^H \pi_h (1 - \pi_h), \quad (2.11)$$

kde π_i apriorní pravděpodobnost.

*Existuje-li například závislost mezi třídícím kritériem a skupinou proměnných, pak je použito užitečné diskriminační kritérium. (pro dvě stejné velké skupiny je to například 0,5)*²⁸

Tato pravděpodobnost se dá zjistit různými způsoby. Nejjednodušší a nejznámější je metoda tzv. resubstituce, kde je diskriminační kritérium použito k třídění těch jednotek, na jejichž základě bylo získáno. Nevýhodou tohoto postupu je, že odhadovanou pravděpodobnost podhodnocuje. Pro rozsáhlé soubory je třeba rozdělit disponibilní datový soubor na dvě části a na základě jedné části odvodit diskriminační kritérium a údaje z druhé části řeknou, jak dobře jsou jednotky s užitím kritéria klasifikovány. Tento postup se dá ale použít pouze na rozsáhlé soubory. Další typem vyhodnocování je například tzv. křížová kontrola (jackknife) správnosti diskriminace.²⁹

²⁸ Citace HEBÁK, P. -- JAROŠOVÁ, E. -- PECÁKOVÁ, I. -- PLAŠIL, M. -- ŘEZANKOVÁ, H. -- VILIKUS, O. -- VLACH, P. *Statistické myšlení a nástroje analýzy dat 2013 s 207*

²⁹ PECÁKOVÁ, I. *Statistika v terénních průzkumech*. 2011 s 198

2.4 Metoda klasifikačních a regresních stromů

Rozhodovací stromy jsou jednou z nejvyužívanějších metod v oblasti klasifikačních úloh učení s učitelem, konkrétně v data miningu. Metoda klasifikačních a regresních stromů (CART) patří do tohoto odvětví.³⁰

2.4.1 Základní algoritmus rozhodovacího stromu

Rozhodovací stromy patří k nejznámějším algoritmům z oblasti metody strojového učení. Při tvorbě rozhodovacích stromů se postupuje na základě metody „rozděl a panuj“ respektive divide and conquer. To znamená, že zvolená data se postupně dělí na stále menší podmnožiny (uzly stromu) tak, aby v těchto podmnožinách převládaly příklady jedné třídy. Klasický postup při tvorbě rozhodovacího stromu, kdy data na začátku tvoří jednu množinu a na konci jsou vzniklé podmnožiny obsahující příklady téže třídy, se nazývá Top down induction of decision trees (TDIDT). Začíná se tedy jedním uzlem takzvaným kořenem, od kterého je postup směrem dolů, a tím je tvořen rozhodovací strom.³¹

Obecný algoritmus rozhodovacích stromů by se dal rozdělit do tří kroků:

- 1) Zvolení jeden atribut jako kořen dílčího stromu.
- 2) Rozdělení dat v tomto uzlu na podmnožiny podle daných hodnot zvoleného atributu a přidání uzlu pro každou podmnožinu.
- 3) Existuje-li uzel, pro který nepatří všechna data do téže třídy, pak je pro tento uzel opakován postup od bodu I, jinak konec procesu.

Nejdůležitější na celém algoritmu je určitě otázka, jak vybrat vhodný atribut pro větvení stromu (bod I). Je nezbytné zvolit takový atribut, který od sebe nejlépe odliší příklady různých tříd. Tato volba je prováděna pomocí charakteristiky atributu převzaté z teorie informace a pravděpodobnosti: entropie, informační zisk, poměrný informační zisk, χ^2 , Gini index a další. V tomto případě, jestliže bude použita metoda CART, je

³⁰ BERKA, P. *Dobývání znalostí z databází*. 2003 s 86-87

³¹ BERKA, P. *Dobývání znalostí z databází*. 2003 s 86 a Kingsford, C., & Salzberg, S. L. (2008). What are decision trees? *Nature Biotechnology* [online]

podstatný Gini index, který je popsán v samostatné podkapitole. Nyní je zde jako příklad uvedena Entropie.³²

Profesor Petr Berka (2003) píše: „Entropie vyjadřuje míru neuspořádanosti nějakého systému. V teorii informace je entropie definována jako funkce

$$H = \sum_{t=1}^T (p_t \log_2 p_t), \quad (2.12)$$

kde p_t je pravděpodobnost výskytu t -té třídy (v našem případě relativní četnost třídy t počítaná na určité množství příkladů) a T je počet tříd. Výpočet entropie pro jeden atribut se provádí tímto způsobem; pro každou hodnotu v , kterou může nabýt uvažovaný atribut A , spočítáme podle uvedeného vzorce entropii $H(A(v))$ ve skupině příkladů, které jsou pokryty kategorií $A(v)$

$$H(A(v)) = - \sum_{t=1}^T \frac{n_t(A(v))}{n(A(v))} \log_2 \frac{n_t(A(v))}{n(A(v))}. \quad (2.13)$$

Spočítáme střední entropii $H(A)$, jako vážený součet entropií $H(A(v))$, přičemž váhy v součtu jsou relativní četnosti kategorií $A(v)$ v datech D_{TR}

$$H(A) = - \sum_{v \in \text{Val}(A)} \frac{n(A(v))}{n} H(A(v)). \quad (2.14)$$

Pro větvení stromu pak vybereme atribut s nejmenší entropií $H(A)$.,³³

2.4.2 Prořezávání stromů

Proces rozhodovacích stromů končí až ve chvíli, kdy všechny příklady odpovídající jednotlivým listovým uzlům patří do stejné třídy. Někdy ovšem tento postup není možný a také ani žádoucí. Bezchybná klasifikace trénovacích dat není možná v situacích, kdy jsou trénovací data zatížena šumem. Výsledný strom se může stát příliš košatým a nepřehledným a požadavek na bezchybnou klasifikaci trénovacích dat může vést

³² BERKA, P. *Dobývání znalostí z databází*. 2003 s 86-87

³³ Citace BERKA, P. *Dobývání znalostí z databází*. 2003 s 87

k přeučení (overfitting). Proto se často uvádí požadavek, aby v listovém uzlu převažovaly příklady jedné třídy. Tato změna změní výsledný strom a to tak, že je menší, srozumitelnější pro interpretaci, ovšem za cenu zhoršeného chování při klasifikaci těchto dat.³⁴

Profesor Berka (2003) ve své knize dodává: „K tomuto upravenému (redukovanému) stromu se lze propracovat 2. způsoby:

- a) Modifikací původního algoritmu (redukovaný strom vytvoříme přímo).
- b) Následným prořezáním stromu (post-pruning úplného stromu).

V praktických úlohách se spíše používá druhý způsob převážně z důvodu, že je obtížné poznat, kdy předčasně ukončit růst stromu. Při vytváření redukovaného stromu (ať už prvním nebo druhým způsobem) je základní otázkou, jak poznat, kdy lze nelistový uzel nahradit listem. K tomu můžeme použít buď nová data (tzv. validační), která se použijí pro testování uvažované redukce, nebo odhadneme vhodnost redukce na základě statistického testu pouze z trénovacích dat.“³⁵

Obecný algoritmus prořezávání lze tedy zapsat následovně:³⁶

- 1) Převedení stromu na pravidla.
- 2) Generalizace pravidla o odstranění podmínky z předpokladu, pokud dojde ke zlepšení odhadované správnosti.
- 3) Uspořádání prořezávaných pravidel podle odhadované správnosti. V tomto pořadí jsou pak pravidla použita pro klasifikaci.

2.4.3 Regresní stromy

Profesor Berka (2003) popisuje regresní stromy ve své knížce: „Rozhodovací stromy se dají rozdělit do dvou skupin. První skupina, kdy jsou vytvořeny stromy pro klasifikaci určitých objektů do tříd, nazýváme klasifikační stromy. Za pomoci regresních stromů je možné naopak odhadovat hodnotu nějakého číselného atributu. V listových uzlech mají

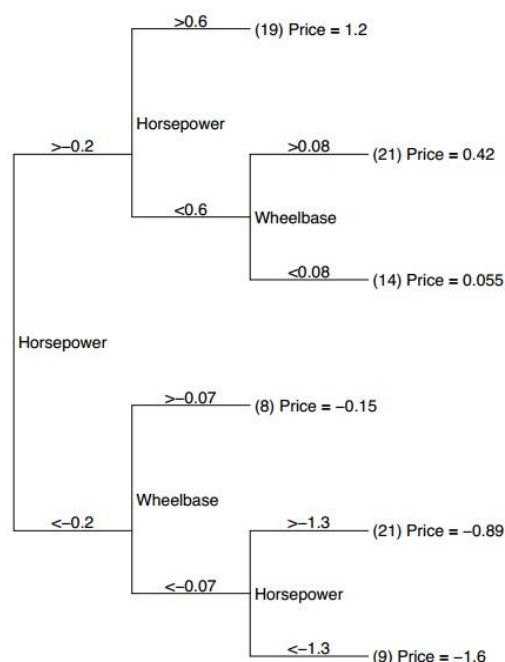
³⁴ BERKA, P. *Dobývání znalostí z databází*. 2003 s 94-95

³⁵ Citace BERKA, P. *Dobývání znalostí z databází*. 2003 s 94-95

³⁶ BERKA, P. *Dobývání znalostí z databází* 2003 s 95

takové stromy místo názvu třídy například konkrétní hodnotu (konstantu), která odpovídá průměrné hodnotě cílového atributu.³⁷

Obrázek 2: Příklad regresního stromu



Zdroj: SPEYBROECK, N. (2012) [Online]

A vzápětí dodává: „Algoritmus pro tvorbu regresního stromu je obdobný jako algoritmus TDIDT. Rozdíl spočívá ve volbě atributu pro větvení. Vycházeno je ze směrodatné odchylky hodnot cílového atributu.

$$S_y = \sum_{v \in \text{Val}(A)} \frac{n(A(v))}{n} S_y(A(v)); \quad (2.15)$$

kde S_y^2 je rozptyl hodnot cílového atributu pro celá data a $S_y^2(A(v))$ značí rozptyl hodnot cílového atributu pro příklady pokryté kategorií $A(v)$. Pro větvení je vybrán atribut,

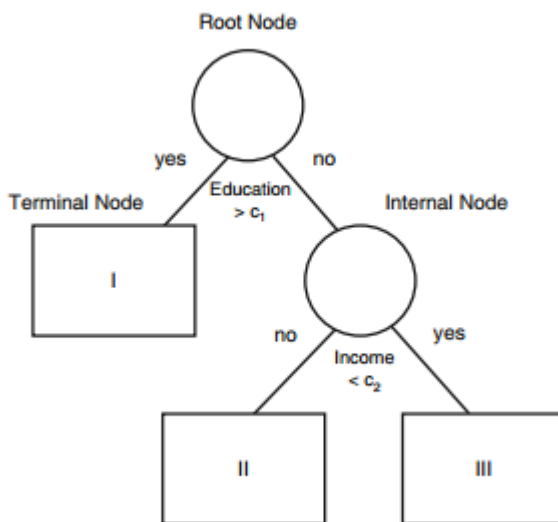
³⁷ Citace BERKA, P. *Dobývání znalostí z databází*. 2003 s 99

který maximalizuje toto kritérium. Větvení skončí, pokud se hodnota cílového atributu pro příklady v uvažovaném uzlu jen málo liší.³⁸

2.4.4 Metoda CART

Metoda CART (Breiman, Friedman, Olshen, Stone, 1984) neboli classification and regression trees využívá binární přístupu k rozhodovacím stromům, kdy je pokládána jednoduchá otázka a odpověď na ni určí, jaká otázka bude následovat. V každém nelistovém uzlu se tedy data rozdělují do dvou množin. Jako koncové listy jsou brány otázky, na které už nemohou být kladeny další.

Obrázek 3: Příklad rozhodovacího stromu metody CART



Zdroj: SPEYBROECK, N. (2012) [Online]

Tato metoda na rozdíl od metod, které využívají entropii, využívá Giniho index a její největší výhodou je, že dokáže krom klasifikačních stromů vytvářet také i stromy regresní.³⁹

³⁸ Citace BERKA, P. *Dobývání znalostí z databází*. Praha: Academia, 2003 s 99

³⁹ BERKA, P. *Dobývání znalostí z databází*. 2003 s 100 a Speybroeck, N. (2012). Classification and regression trees. *International Journal of Public Health* ,[online]

2.4.4.1 Algoritmus metody CART

Jak je naznačeno výše, metoda CART využívá místo entropie Giniho index. Tento index se spočítá jako

$$Gini = 1 - \sum_{t=1}^T (p_t^2), \quad (2.16)$$

kde p_t je opět relativní počet příkladů t -té třídy zjišťovaný na nějaké (pod)množině. Hodnota indexu je minimální v případě, že příklady patří do jedné ze tříd, a maximální v případě, že příklady jsou rovnoměrně rozděleny mezi obě třídy.⁴⁰

Profesor Berka ke Giniho indexu píše: „Hodnota Giniho indexu pro jeden konkrétní atribut je spočítána stejně, jako hodnota entropie pro jeden atribut. Tedy tak, že je spočítán vážený součet indexu pro jednotlivé hodnoty atributu, přičemž váhy budou opět relativní četností příslušných hodnot.

$$Gini(A) = \sum_{v \in Val(A)} \frac{n(A(v))}{n} Gini(A(v)) \quad Gini(A(v)) = 1 - \sum_{t=1}^T \left(\frac{n_t(A(v))}{n(A(v))} \right)^2. \quad (2.17)$$

Pro větvení je použit ten atribut, který bude mít nejmenší hodnotu tohoto indexu. Také je možné maximalizovat rozdíl mezi Gini indexem pro celkový atribut a Gini indexem jednoho atributu.⁴¹

2.4.5 Použití rozhodovacích stromů

Rozhodovací stromy patří mezi využívané metody pro klasifikaci dat, protože se u nich vychází z jednoduchých pravidel a také jsou v rámci možností rychlé, měřitelné a konkrétně metoda CART se dá použít i pro velký počet proměnných. Jednou z dalších výhod je, že rozhodovací stromy nejsou ovlivněny multikolinearitou a možnou korelací mezi proměnnými.⁴²

⁴⁰ BERKA, P. *Dobývání znalostí z databází. 2003 s 89*

⁴¹ Citace BERKA, P. *Dobývání znalostí z databází. 2003 s 89-90*

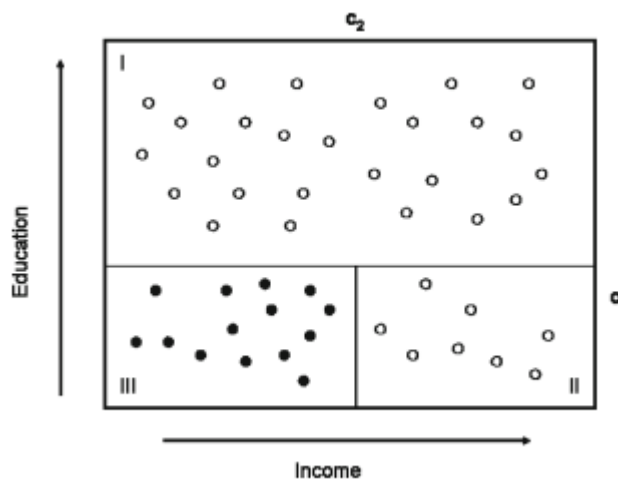
⁴² Berry, M. J. A. - Linoff, G. *Data mining techniques for marketing, sales, and customer relationship management. 2004 s 84*

Rozhodovací stromy jsou vhodné pro úlohy kde:⁴³

- 1) Příklady jsou reprezentovány hodnotami atributů.
- 2) Úkolem je klasifikovat příklady do malého počtu tříd.
- 3) Hledaný popis konceptů může být tvořen disjunkcemi.
- 4) Trénovací data mohou obsahovat chybějící hodnoty.

Rozhodovací stromy rozdělují prostor daných atributů nadrovinami rovnoběžnými s osami souřadné soustavy, jak je vidět na následujícím obrázku č. 4.

Obrázek 4: Příklad vyjadřovací síly rozhodovacích stromů



Zdroj: SPEYBROECK, N. (2012) [Online]

⁴³ BERKA, P. *Dobývání znalostí z databází. 2003 s 101-102 a Speybroeck, N. (2012). Classification and regression trees. International Journal of Public Health, [online]*

3 Teoretická východiska

Sociologové, podnikatelé, ekonomové, politici a běžní občané všeobecně vnímají slovo „modernizace“ jako určitý proces, pomocí něhož se uvádí v podniku, ve strojírenství, IT či jiném odvětví daný proces, technologie nebo strategie na současný stav či trend. Typickým příkladem může být například továrna, kde se zavádí nová technologie či nové, respektive modernější stroje do výroby. Co ovšem přímo souvisí se slovem modernizace, jsou modernizační teorie. Tyto teorie popisují progresivní proces, pomocí něhož je společnost převáděna z pre-moderního na tradiční a poté na moderní typ společnosti.⁴⁴

Podle názorů odborníků by měl být modernizován veřejný sektor, trh práce, společnost zaměstnání, ale také celý sociální stát. Procesem modernizace by měl také projít životní styl, hodnoty a postoje obyvatel. Tyto názory ovšem nešíří pouze liberální politici, ale také zástupci konzervativní politické části, kteří také na naprostou nutnost modernizovat upozorňují. Slovíčko modernizace se stalo samo o sobě jakýmsi odzbrojujícím faktorem při debatování o nejrůznějších problémech. Stačí například říci, že se jedná o modernizační reformu, a lidé automaticky předpokládají, že se jedná o změnu k lepšímu. Ani nyní po téměř půlstoletí debat nejrůznějších sociologů ale není jednota názoru, co to slovo „modernizace“ ohledně populace vlastně znamená.⁴⁵

Encyklopedie Britannica popisuje modernizační teorie jako úvahy, pomocí kterých se snaží nalézt proměnné, jež přispívají k posunu tradiční společnosti k moderní. Britannica dále na základě děl několika významných sociologů tvrdí, že postupem času dojde k další modernizaci společnosti na úplně nový typ. Zastánci teorie modernizace tvrdí, že státy, které tímto procesem prošly, jsou bohatší, silnější a jejich občané si mohou dopřát novější technologie a výrobky „moderní“ doby, zatímco její odpůrci namítají, že tyto procesy sice vedou k technologickému pokroku, ale také ničí mezilidské vztahy.⁴⁶

O rozlišení společnosti pomocí psychologických rysů se pokouší sociolog Anthony Giddens, který při srovnávání předmoderní a moderní společnosti rozlišuje prostředí důvěry a prostředí rizika. U předmoderní společnosti uvádí v prostředí důvěry tyto aspekty: příbuzenské vztahy, místní komunitu, náboženskou kosmologii a tradici. Naopak v

⁴⁴ KNOBL, W. "Theories That Won't Pass Away: The Never-ending Story." in Gerard Delanty, Engin F. Isin, eds. *Handbook of Historical Sociology* (2003): 96-107 [online]

⁴⁵ KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 9

⁴⁶ KUMAR, K. *Modernization*. Encyclopedia Britannica. Britannica.com. Retrieved on 2013-08-17 [online]

moderní společnosti jsou tu aspekty vztahů osobního přátelství, abstraktních systémů, kterými Giddens (2010) myslí prostředky pro stabilizaci vztahů, a nakonec uvádí aspekt na budoucnost orientovaného myšlení. Prostředí rizika u předmoderní společnosti znamená hrozby ze strany přírody, hrozbu lidského násilí a riziko ztráty boží milosti. U moderní společnosti jsou zástupci prostředí rizika hrozby vzešlé z reflexivity modernity, hrozba lidského násilí z hlediska industrializace války a hrozba ze ztráty osobního smyslu. Na první pohled je zde jasný rozdíl. Zatímco předmoderní společnost je více nábožensky založená a spíše zaměřená na tradice a rodinu, moderní společnost myslí neustále do budoucna a spíše se bojí toho, jak na ni bude nahlíženo s odstupem času.⁴⁷

Giddens (2010) také uvádí jakýsi problém modernizace, kterým je ztráta kvalifikace a rekvalifikace populace v každodenním životě. Mnoho sociologů se podle něj domnívá, že na rozdíl od předešlé předmoderní doby je dnes v podstatě všechno jasné a způsob jak „svět funguje“ může být vyčerpávajícím způsobem popsán. To ovšem podle něj není pravda. Jako příklad uvádí, že nyní například málokdo, když vypíná či zapíná světlo, ví, odkud se daná elektřina bere a jak vzniká.⁴⁸

Podle sociologa Jana Kellera (2007) mají nejasnosti ohledně obsahu modernizace tři hlavní příčiny. Za první příčinu považuje to, že teorie modernizace nejsou vyústěním určitého vnitřního procesu poznání. Další příčinu vidí v tom, že se změnil objekt modernizačního úsilí, protože v polovině 20. století šlo o rozšíření vlivu Spojených států amerických na země světového jihu, koncem téhož století se jedná o transformaci zemí Východu a o jejich integraci do západního modelu, jehož nejmodernější variantu údajně opět představují Spojené státy. A poslední tedy třetí faktor se týká samotných západních zemí. V 20. století se nijak výrazně nezměnila západní společnost, ke které by měly ostatní země vzhlížet a podle jejich obrazu být formovány.⁴⁹

Ve skutečnosti dnes svět zažívá již druhou vlnu modernizace. První vlna proběhla krátce po druhé světové válce a vyvrcholila v padesátých a počátkem šedesátých let. Následně ztratila na vlivu a na dalších zhruba dvacet let upadly modernizační teorie téměř v zapomnění. Teprve v osmdesátých letech proběhla jakási renesance modernizačních

⁴⁷ GIDDENS, A. a MÜLLER A.. *Důsledky modernity / Anthony Giddens ; [z anglického originálu .. přeložil Karel Müller]*. 2010. ISBN 9788074190353 s 93

⁴⁸ GIDDENS, A. a MÜLLER A.. *Důsledky modernity / Anthony Giddens ; [z anglického originálu .. přeložil Karel Müller]*. 2010. ISBN 9788074190353 s 128-130

⁴⁹ KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 9-10

teorií a jedná se o trend, který přetrvává dodnes. Obě tyto „vlny“ jsou ve zkratce popsány v této práci.⁵⁰

3.1 První vlna modernizace

Pro začátek je nejprve dobré ujasnit si rozdíl mezi modernizací a modernitou. O to se pokouší sociolog Jan Keller (2007): „*Jestliže chceme posoudit obsah, přínos a nedostatky modernizačních teorií, musíme si nejprve ujasnit rozdíl mezi teorií modernizace a teorií modernity. Teorie modernity představují velké a poměrně abstraktní myšlenkové systémy, s jejichž pomocí se klasičtí sociologové 19. století a po nich řada teoretiků této disciplíny ve 20. století snažili jednak objasnit povahu moderní společnosti v protikladu k tradičním společnostem, jednak pojmenovat síly, které stály u vzniku modernity, a také popsat procesy, skrze které se tato nová a zcela ojedinělá fáze v lidských dějinách zformovala. Teorie modernity se tedy snaží zpětně vysvětlit, co se v moderní době přihodilo se společností, jaké změny v ní proběhly a stále probíhají.*“⁵¹

Anthony Giddens (2010) rozlišuje tři dominantní zdroje dynamického vývoje modernity. První zdroj je podle něj oddělení času a prostoru ten má za následky to, že se dají časové a prostorové parametry přesně oddělit do určitých zón. Druhý zdroj je rozvinutí vyvazujících mechanismů. Tento mechanismus reorganizuje sociální vztahy napříč velkými vzdálenostmi a časem a vytlačuje tradiční sociální činnosti. Třetí zdroj je podle Giddense reflexivní přivlastňování vědění. Giddens zde myslí to, že jakési systematické vědění o životě se stává integrální částí a vytlačuje tak sociální život z ustálené tradice. Podle Giddense (2010) spolu všechny tři části souvisí.⁵²

Německý sociolog Wolfgang Knöbl dokumentuje, že modernizační teorie se z hlediska vývoje sociologického myšlení objevily prakticky z ničeho nic. Především v předchozím vývoji americké sociologie nebylo nic, na co by tyto teorie mohly navázat nebo od čeho se odrazit. Usuzuje tak především z toho důvodu, že americká sociologie po

⁵⁰ Citace KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 9

⁵¹ Citace KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 18

⁵² GIDDENS, A. a MÜLLER A.. *Důsledky modernity / Anthony Giddens ; [z anglického originálu .. přeložil Karel Müller]*. 2010. ISBN 9788074190353 s 52

druhé světové válce nedisponovala žádnou makroteorií, která by dokázala zachytit proces historické změny, což je považováno za jádro teorie modernizace.⁵³

Jednu z prvních modernizačních teorií se pokusil vytvořit americký sociolog Marion Levy. Levy vycházel především z poznatků dalšího amerického teoretika Talcotta Parsonsa, který počátkem třicátých let dvacátého století pomocí své analýzy profesního života nastínil určité „vzorce proměnných“, které poté Levy využil při pokusu uchopit problémy při přechodu od tradiční k moderní společnosti.⁵⁴

Jan Keller (2007) k Levyho pokusu píše: „*Teorie modernizace díky Levyho práci získala jakýsi balíček politických a společenských předsudků. Tyto předsudky vznikly jako zobecnění životního stylu a uvažování příslušníků různých vrstev. Hodnoty těchto vrstev byly následně propojeny s evolučním učením a s abstraktní teorií strukturního funkcionalismu.*⁵⁵ *Výsledné schéma bylo velmi prosté: Tvůrcem prosperity, štěstí a pokroku je průmyslová společnost dosahující vysokého tempa růstu. Tato společnost je nesena duchem podnikavosti a osobní výkonnosti. Takový postoj ke světu je neslučitelný s rutinou, lenivostí a omezeními, jež vládou v tradiční společnosti. Je třeba, aby v zaostalých společnostech převládlo nové, čínorodé myšlení a podnikavý duch. To nastartuje ekonomický růst a v nastalém blahobytu se bude dobře dařit demokracii.*“⁵⁶

V polovině 60. let 20. století ale přišlo poněkud velké vystřízlivění z amerického snu. V amerických velkoměstech se začala masově objevovat bída, začátek boje černochů za lidská práva, šok z válečných zločinů spáchaných americkou armádou ve Vietnamu a to ukázalo, že tolik obdivovaná poválečná prosperita byla pouze výsledkem mimořádné ekonomické, politické a sociální situace.⁵⁷

Další americký autor F. W. Riggs ve své práci, která byla vydaná o pár let později než práce Rostowa, tvrdí, že ekonomické a sociální aspekty společnosti mají takovou povahu, že je téměř nemožné reprodukovat v jejich rámci instituce mechanicky okopírované ze západních zemí. V rozvojových zemích díky tomu dochází k rozkladu tradiční společnosti, aniž by se modernizovaly instituce a chování společnosti. Tyto

⁵³ KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 19

⁵⁴ KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 20

⁵⁵ KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 21

⁵⁶ Citace KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 21

⁵⁷ KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 23-24

neočekávané výsledky modernizace v rozvojových zemích vedly ke zpochybnění vymezení výrazů tradice a modernita, které byly do té doby jednoznačně definovány.⁵⁸

Od této doby začínají někteří sociologové, například izraelský Samuel N. Eisenstadt, zpochybňovat do té doby jasně definované schéma modernizačních teorií. Eisenstadt uvádí, že státy, které se vyvíjely jinak než Západ, nutně nemusí být ponořeny do tradice, ale jednoduše mohou být „naprogramovány“ odlišně než západní země.⁵⁹

V polovině šedesátých let, kdy západní státy prožívaly poněkud problematický poválečný vývoj, byly vážně zproblematizovány všechny původní předpoklady teorie modernizace a jejich velká kritika měla za následek to, že autoři těchto teorií se začali vyjadřovat čím dál abstraktněji. Pod náporom této kritiky první vlna modernizačních teorií pomalu ale jistě slábla a téměř na dvacet upadla v zapomnění.⁶⁰

3.2 Druhá vlna modernizace

Sociolog Jan Keller (2007) tvrdí: „*Druhá vlna modernizace nemá ani zdaleka relativní jednoduchost vlny první. Prolínají se v ní totiž dva myšlenkově značně odlišné zdroje a dva radikálně různé pohledy na modernitu, jejichž polemické střetávání probíhá již od poloviny sedmdesátých let a s postupem času neustále sílí.*“⁶¹

Prvním proudem je jednorozměrná teorie modernizace. Tato teorie je pokračovatelem první vlny modernizace, která je jen nepatrně pozměněna, ale v základě uznává všechny základní teorie vyslovené již v 50. letech 20. století. Od 50. let byl tento myšlenkový proud na dvě desetiletí v nezájmu a až počátkem let osmdesátých se znovu objevil, a to především díky ekonomickému vzestupu asijských národů.⁶²

Druhým typem jsou emancipační teorie. Tyto teorie vznikají jako reakce na technologické, ekologické, sociální a politické problémy, které se od šedesátých let 20. století objevují uvnitř samotné moderní společnosti, a to s rostoucí intenzitou. Dalo by se říci, že jsou jistým výrazem nespokojenosti s jednorozměrnými teoriemi modernizace,

⁵⁸ KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 25-26

⁵⁹ KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 27

⁶⁰ KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 28

⁶¹ Citace KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 34

⁶² KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 34

kteře mají tendenci všechny tyto problémy znevažovat, a hlavně je řešit podle postupů, jež byly sepsány ještě dřívě, že se tyto problémy naplno objevily.⁶³

V druhé vlně modernizačních teorií je zajímavá určitá absurdita situace, kdy teorie, které měly v první vlně znamenat šanci pro rozvojové země, aby se postupem času staly stejně vyspělými a „moderními“, jsou nyní samy modernizovány tak, aby posílily šance, že dané rozvojové země budou schopny překonat problémy plynoucí z jejich vlastního vývoje. Můžeme tak sledovat to, že teorie, které byly k dispozici ke konci 20. století vyspělým zemím při řešení jistých problémů, byly v polovině téhož století doporučovány pro rozvoj zemí třetího světa.⁶⁴

Nespokojenost s jednorozměrnými teoriemi plynula především z toho, že teorie se v podstatě téměř vůbec nezměnily, ale život členů moderní společnosti mezitím doznal výrazných změn. Přechod k postindustriální společnosti měl za následek nejistotu zaměstnanců, a to souvisí s úbytkem velkých formálních organizací, se zvýšením tlaku na globální konkurenceschopnost a další problémy. Kvůli těmto situacím je plánování do budoucnosti velmi ztíženo a možnost jaksi ovládat okolnosti svého života se tím pádem stává nesplnitelným cílem.⁶⁵

Dílo, které dobře vystihuje tuto situaci, napsal sociolog Wolfgang Zapf, který se problému modernizace věnoval už od sedmdesátých let 20. století. Tento autor se snaží pomoci rozvojovým zemím dostat se na úroveň západních zemí, i když připouští, že první vlna modernizace nedopadla příliš úspěšně. Zapf se snaží najít příčiny, které vedly k tomu, že první vlna modernizace nebyla úspěšná. Zjišťuje, že hlavním problémem bylo to, že modernizátoři se snažili rozvojovým zemím předem naordinovat postupy, podle kterých měly země postupovat, ať se děje cokoli, a přitom nevzali v potaz důsledky, které mohou skrz danou modernizaci pro zemi nastat. Zapf se této problematice věnuje až do 90. let.⁶⁶

Sociolog Jan Keller (2007) píše: „*V případě otevřené modernizace vyspělých zemí, kde cíl není předem dán, je tím jediným, co je pevně dáno, institucionalizovaný prostor, v jehož rámci se inovace mohou pohybovat. Jde o instituce trhu, stranické demokracie, sociálního státu a masového konzumu. Stále znovu se tak vrací stejná myšlenková figura.*

⁶³ KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 34

⁶⁴ KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 38

⁶⁵ KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 38

⁶⁶ KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 39

Wolfgang Zapf připouští, že všechny prvky, které považuje za klíčové pro modernizaci, procházejí závažnou krizí. Hovoří o krizi, která zastavila poválečný růst, o krizi masového konzumu, o krizi sociálního státu, o krizi legitimacy politiky. Všechny tyto krize mají být překonány inovacemi, které se pohybují v rámci ekonomiky růstu, masové spotřeby a sociálního státu. Neustálý proud inovací tak má udržet vývoj v rámci mantinelů, které vytyčila první vlna modernizace, a to bez ohledu na to, že tato vlna nesplnila to, co se od ní očekávalo.“⁶⁷

Zapfovo pojetí je tedy podobné tomu, které vzniklo v 50. letech s tím, že je zde doplněno vědomí akutních hrozeb, jak tomu bylo v 70. letech. V podstatě nám říká, že čím silnější krize lidstvo postihne, tím je to pro lidstvo lepší, protože přináší silnější impulz k inovacím všeho druhu.⁶⁸

⁶⁷ Citace KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 43-44

⁶⁸ KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 44

3.3 Globalizace

Jestli lze něco označit jako přirozené vyústění procesu modernizace, je to právě globalizace. Ač spolu tyto dva na první pohled rozdílné procesy nemají příliš společného, pravdou je, že proces globalizace je přímým výsledkem procesu modernizace. Globalizace jako taková ovšem nezasahuje pouze do makroprocesu, ale také do života běžného obyvatele dané země. S globalizací totiž zanikají hranice a přichází velké nadnárodní společnosti, které existují i bez státu a bez světové vlády. Sociolog Jan Keller tvrdí, že globalizace není ničím jiným než pouze celosvětovým pokusem o kolonizaci veřejného prostoru soukromým zájmem.⁶⁹

Anthony Giddens (2010) píše: „*Globalizace může být definována jako zintenzivnění celosvětových sociálních vztahů, které spojují vzdálené lokality takovým způsobem, že místní události jsou formovány událostmi dějícími se mnoho mil daleko a naopak.*“⁷⁰

Jako příklad uvádí, že kdokoliv, kdo v současné chvíli zkoumá v určité části světa města, si je vědom toho, že to, co se děje v jeho okruhu, je ovlivňováno aspekty, které působí i z velké vzdálenosti od tohoto místa, například peněžní trhy.⁷¹

3.3.1 První fáze globalizace

První fáze globalizace začíná v 50. a 60. letech 20. století a má podobu pomoci rozvojovým zemím z jihu planety. Důvodem bylo, že zde jen 15 % obyvatel žilo ve větších městech a zbytek obyvatelstva byl různě „rozmíchán“ napříč zeměmi. Tato rozvojová pomoc byla ale od začátku poskytována špatně, kdy veškeré investice poskytnuté dané zemi sebralo větší město a do okolí, kde byla většina obyvatel, se z financí a modernizace nedostalo téměř nic. Tato událost měla za následek, že průmyslový rozvoj velkých měst, který byl spojen s modernizací, ničil ekonomickou strukturu venkova, který měl za následek masivní migraci z venkova do velkých měst. Výsledkem bylo, že šťastnější část

⁶⁹ KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 121

⁷⁰ Citace GIDDENS, A. a MÜLLER A.. *Důsledky modernity / Anthony Giddens ; [z anglického originálu .. přeložil Karel Müller]*. 2010. ISBN 9788074190353 s 62

⁷¹ GIDDENS, A. a MÜLLER A.. *Důsledky modernity / Anthony Giddens ; [z anglického originálu .. přeložil Karel Müller]*. 2010. ISBN 9788074190353 s 63

menšiny si trochu polepšila, ale větší část byla stále bezmocnější a jejich krize se stále prohlubovala.⁷²

Podle sociologa E. F. Schumachera spočíval hlavní problém této pomoci rozvojovým zemím v tom, že nezvyšovala soběstačnost těm, kterým měla pomáhat, ale naopak pomáhala rozvinout závislost na výrobě, konzumu a práci ve velkých nadnárodních společnostech. Tímto způsobem pak chudé obyvatele nutila do závislosti na bohatých zemích a znemožňovala jim jakoukoliv jinou alternativu.⁷³

V 50. letech země z jihu planety přemýšlely, jestli se vydají strategií zaměření industrializace na vlastní spotřebu, která by snížila závislost na surovinách okolních zemí. Nakonec se, za přispění velkých nadnárodních společností, vydaly přesně opačným směrem. Tyto investice byly pochopitelně velmi finančně náročné, a tak není překvapením, že měly za následek zadlužování těchto zemí. Míra zadlužení vedla následně k tomu, že rozvojové země naprosto otevřely trh se surovinami pro země vyspělé, pro které byl nákup těchto surovin lacinou záležitostí. Zadlužení rozvojových zemí bylo tedy způsobeno velkými půjčkami a není žádným překvapením, že velkou část těchto částek „sebrali“ do kapsy úředníci a politici. Další část půjček byla poskytnuta armádním složkám pro případ, že by místní obyvatelé nesouhlasili s nastoleným trendem.⁷⁴

Jako další důvod globalizace a masivního exodu rozvojových zemí se dají považovat válečné konflikty. Od skončení druhé světové války vzniklo po celém světě na 200 konfliktů, z nichž téměř všechny se odehrály v zemích třetího světa. Demografické příčiny migrace nejsou způsobeny jen přelidněním, ale také tím, že rozvojová pomoc nepomáhá příliš efektivně a neumožňuje chudším obyvatelům příliš hospodařit a získat dostatek obživy v místech, kde žili jejich předkové.⁷⁵

Nejsmutnější příčinou migrace je samotná rozvojová pomoc, kdy jsou finance napumpovány do nesmyslných projektů dálnic, přehrad, kácení dřeva a stavby rozsáhlých areálů a vyhánějí původní obyvatele z jejich obydlí a přirozeného prostředí. Toky těchto

⁷² KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 131

⁷³ KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 131-132

⁷⁴ KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 133

⁷⁵ KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 137

uprchlíků ale nemusí skončit pouze ve velkých městech a mohou se severní společnosti, která byla impulsem těchto změn a investicí, vrátit jako bumerang.⁷⁶

3.3.2 Druhá fáze globalizace

Jan Keller (2007) píše: „*Druhá fáze globalizace začíná obdobně jako konec první fáze. Nadnárodní společnosti od sedmdesátých let stále získávají větší sílu a vliv na řízení států třetího světa. Hlavní výhodou těchto společností je působení ve více státech a bez větších problémů svoje části přesunovat a využívat rozdílnou cenu práce v těchto zemích, protože čím více je firma pohyblivější, tím více může vytvářet tlak na státy, aby absorbovaly co největší část jejich sociálních, environmentálních a dalších nákladů.*“⁷⁷

Toto snadné manévrování mělo za následek to, že v 90. letech mělo deset největších nadnárodních firem větší obrat než hrubý domácí produkt stovky menších států. Posilování vlivu nadnárodních společností mělo vliv i na jejich vnitřní reorganizaci, a tak zatímco navenek jejich vliv čím dál víc sílil, uvnitř se byly schopny zbavovat i svých vlastních zaměstnanců.⁷⁸

Společnosti začaly svoje zdroje co nejvíce soustřeďovat do rozvojových zemí třetího světa. Pracovní podmínky se zde vyznačovaly a stále vyznačují velmi dlouhou pracovní dobou, například i 16 hodin denně, dále pak velmi nízkými mzdami, špatnými hygienickými podmínkami a žádnou činností odborů. V současné době se už některé firmy snaží od tohoto trendu ustupovat a podepisovat různé dokumenty, kde se zavazují k tomu, že tento typ pracovní síly už nevyužívají, ale v době druhé vlny globalizace se jednalo o běžnou věc. V těchto továrnách tak muži, ženy i děti za minimální mzdu vyráběli obuv, hračky, ale také elektroniku a součástky do automobilů. S těmito osobami firmy uzavíraly pracovní vztah na neurčitou dobu, za minimální mzdu a možnosti okamžitého ukončení pracovního poměru.⁷⁹

Šárka Brychtová (2013) ve svém akademickém článku tvrdí, že jakousi druhou rukou globalizace se stala deregulace neoliberální ekonomiky a politiky. Podle ní se stát

⁷⁶ KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 138-139

⁷⁷ Citace KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 139-140

⁷⁸ KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 140

⁷⁹ KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 143-144

vzdává určitých rozhodovacích pravomocí, tím ale pouze snižuje svůj vlastní vliv na přerozdělování zdrojů, a naopak soustřeďuje čím dál více moci do soukromého sektoru.⁸⁰

V polovině 70.let nastala v oblasti migrace nová situace. Zatímco proud migrantů stále sílil, hospodářská recese v kombinaci s ropnou krizí snižovala tempo růstu ve vyspělých zemích. Nadnárodní společnosti se následkem toho snažily odvést pracovní sílu do méně vyspělých zemí, a i když se vyspělé země snažily bránit přistěhovaleckým vlnám, podíly přistěhovalců a uprchlíků v těchto zemích rostl. V tomto období také narostla nelegální imigrace spojená s organizovaným zločinem.⁸¹

Začínalo se zde objevovat i jakési dilema, kdy pro nadnárodní společnost masová migrace nepředstavovala ani ten nejmenší problém, ale pro sociální stát se jednalo o problém velmi zásadní, protože do vyspělých zemí najednou mířili lidé s minimem vzdělání, schopnostmi a lidé z odlišné kultury. Tato migrace byla a je způsobena tím, že neexistuje žádná zákonná odpovědnost nadnárodních společností, aby se o tyto obyvatele postaraly.⁸²

Neschopnost sociálního státu zareagovat na dopady první globalizační vlny má za následek masové uprchlické vlny a problém s jejich zakomponováním do normální společnosti daného státu. S tímto problémem je potom spojena i kriminalita či nárůst extremismu v dané zemi. Tato fáze probíhá i dnes.⁸³

Šárka Brychtová (2013) tvrdí, že ačkoliv se nesnaží jakkoli zpochybňovat potřebu společenského pokroku a potřeby modernizovat, nebude-li u chudých zemí snížen masivní nárůst populace, a naopak u západních zemí omezena spotřeba, může dojít na nejrůznější katastrofy, které mohou mít za následek ohrožení celé populace na planetě. Při těchto tvrzeních se odkazuje na vědce z britské Královské společnosti.⁸⁴

⁸⁰ BRYCHTOVÁ, Š. Globalizace a krize v souvislostech: *Globalization and crisis in connection*. Scientific Papers of the University of Pardubice. Series D, Faculty of Economics & Administration. 20, 29, 5-13, Nov. 2013. ISSN: 1211555X s 6

⁸¹ KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 146

⁸² KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 147

⁸³ KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 s 148

⁸⁴ BRYCHTOVÁ, Š. Globalizace a krize v souvislostech: *Globalization and crisis in connection*. Scientific Papers of the University of Pardubice. Series D, Faculty of Economics & Administration. 20, 29, 5-13, Nov. 2013. ISSN: 1211555X s 10

Jako další negativní dopad globalizace, respektive modernizace uvádí sociolog Ulrich Beck a také encyklopedie Britannica industrializaci produkce. Tato industrializace nezná hranic a výfukové plyny z továren tak unikají do ovzduší, které „nahlodává“ historické památky a sochy. Beck jako příklad nejprve uvádí Kanadu, kde jsou jezera znečištěna kyselými dešti a jako druhý příklad Skandinávii, kde i na severu odumírají lesy.⁸⁵

Jak z výše uvedeného vyplývá, pojem modernizace tedy neznamená pouze jakýsi proces výměny něčeho staršího na aktuální stav. Jedná se o rozsáhlý sociologický proces, který s sebou nese nejrůznější následky, které mají dopad na celkovou populaci. Proces modernizace je pouhým začátkem celého řetězce, na jehož konci může stát masivní globalizace, migrace a narůstající kriminalita či ničení životního prostředí. Takovéto následky si však málokdo uvědomuje.⁸⁶

⁸⁵ BECK, U. *Riziková společnost: na cestě k jiné moderně*. Vyd. 1. Přeložil Otakar VOCHOČ. Praha: Sociologické nakladatelství, 2004. Post. ISBN 80-86429-32-6 s 48 a KUMAR, K. *Modernization*. Encyclopædia Britannica. Britannica.com. Retrieved on 2013-08-17 [online]

⁸⁶ KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3 a BRYCHTOVÁ, Š. Globalizace a krize v souvislostech: *Globalization and crisis in connection*. Scientific Papers of the University of Pardubice. Series D, Faculty of Economics & Administration. 20, 29, 5-13, Nov. 2013. ISSN: 1211555X s 10

A BECK, U. *Riziková společnost: na cestě k jiné moderně*. Vyd. 1. Přeložil Otakar VOCHOČ. Praha: Sociologické nakladatelství, 2004. Post. ISBN 80-86429-32-6 s 48 a KUMAR, K. *Modernization*. Encyclopædia Britannica. Britannica.com. Retrieved on 2013-08-17 [online]

4 Vlastní práce

4.1 Popis datového souboru

Pro diplomovou práci byl zvolen datový soubor s názvem „Veřejnost jako aktér modernizace“. Tento průzkum byl pro akademické účely poskytnut Českým sociálněvědním datovým archivem, zkráceně ČSDA. Průzkum byl proveden společností Factum Invenio. Šetření proběhlo na reprezentativním vzorku 2353 obyvatel od věku 15 let vybraných metodou náhodného výběru. Šetření bylo uskutečněno po celém území České republiky. Respondentům bylo položeno celkem 353 otázek, na které odpovídali.

Výzkum byl realizován na základě stratifikovaného vícestupňového náhodného výběru. V rámci stratifikace bylo celé území České republiky rozděleno na 57 částí – strat, ve kterých byl realizován třístupňový náhodný výběr. Při prvním stupni výběru byly vybrány základní sídelní jednotky, při druhém stupni domácnosti a při třetím stupni budou určeni jednotlivci respondenti. Průzkum byl realizován metodou osobních standardizovaných rozhovorů – Face to face s papírovými dotazníky, to znamená pomocí metody PAPI (Pen and paper interviewing). Průměrná délka rozhovoru byla 60 minut. Za neúspěšné rozhovory byly považovány i ty případy, kdy vybraný respondent byl sice kontaktován, ale rozhovor odmítl.

4.2 Analýza závislosti přeměny společnosti na tempu modernizace

V teoretických východiscích se uvádí, že modernizace v posledních letech nabrala opravdu velké tempo. Na základě toho byla v první části analýzy zvolena jako vysvětlovaná proměnná formulace „*Do jaké míry souhlasíte s následujícími výroky – Nezbyvá nám nic jiného, než se přizpůsobit západnímu světu?*“.

Jako vysvětlující proměnné byly vybrány otázky: „*Do jaké míry souhlasíte s výrokem – Naše společnost se mění tak rychle, že to člověk ani nestačí sledovat?*“, „*Ohlédnete-li se zpět, myslíte si, že změna společenského systému u nás, k níž došlo v roce 1989, stála za to?*“ a „*Vy sám/sama jste ochoten/na v případě potřeby zvýšit své pracovní tempo?*“

První tabulka říká, že v tomto případě bylo odebráno 263 respondentů ve prospěch analýzy. Důvodem odebrání těchto pozorování byla eliminace odlehlých hodnot, které by měly vliv na normalitu rozdělení.

Tabulka 1: Výstup z programu IBM SPSS – Analýza proměnných v analýze č.1

Unweighted Cases		N	Percent
Valid		2090	88,8
Excluded	Missing or out-of-range group codes	263	11,2
	At least one missing discriminating variable	0	,0
	Both missing or out-of-range group codes and at least one missing discriminating variable	0	,0
	Total	263	11,2
Total		2353	100,0

Zdroj: Vlastní práce, ČSDA

V tabulce č. 31, která je uvedena v příloze je uvedeno rozdělení respondentů do jednotlivých možností u vysvětlované otázky. Nejvíce respondentů zvolilo v odpovědi na položenou otázku možnost „spíše souhlasím“. Jen o 9 méně má možnost „spíše nesouhlasím“. Obyvatelé České republiky tedy nejsou v této otázce příliš jednotní a názory jsou velmi rozpolcené.

V následující tabulce č. 2 jsou diskriminační schopnosti jednotlivých rovnic. Podstatný je zde sloupec Eigenvalue, který popisuje, jak velkou diskriminační schopnost mají jednotlivé funkce. Další sloupec je pak procentuální rozdělení diskriminačních schopností. Následující sloupec je pak podíl diskriminačních schopností a poslední sloupec je pak kanonická korelace proměnných a skupin, které obsahuje naše vysvětlovaná proměnná. Největší vypovídající hodnotu má první diskriminační funkce, ale hodnota 8,6 % není příliš vysoká.

Tabulka 2: Výstup z programu IBM SPSS – Wilksovo lambda rozdělení v analýze č.1

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	,086 ^a	94,8	94,8	,281
2	,004 ^a	3,9	98,7	,060
3	,001 ^a	1,3	100,0	,034

Zdroj: Vlastní práce, ČSDA

Na základě tabulky č. 2 je možné předpokládat i slabší výsledky Wilksova lambda rozdělení. Nejhodnotnější výsledek zde má první funkce. Čím menší hodnota u Wilksova lambda rozdělení, tím lepší výsledek pro diskriminační schopnosti. Hodnota ve sloupci Sig. značí, že šance, že je naměřená závislost náhodná, je menší než 5 %.

Tabulka 3: Výstup z programu IBM SPSS – Wilksovo lambda rozdělení v analýze č.1

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 3	,917	181,100	9	,000
2 through 3	,995	9,830	4	,043
3	,999	2,410	1	,121

Zdroj: Vlastní práce, ČSDA

V následující tabulce č. 4 jsou pak již uvedeny hodnoty jednotlivých diskriminačních koeficientů. Na základě předchozích výsledků je patrné, že největší diskriminační schopnost má první rovnice, a proto je možné si tyto koeficienty přepsat ve formě rovnice.

Funkce 1 = 0,797*měnicí společnost – 0,402*změna v roce 89 – 0,389*zvýšení pracovního tempa

Největší závislost vykazuje otázka týkající se toho, jestli se společnost mění tak rychle, že to člověk ani nestačí sledovat. Tento výsledek jednoznačně odpovídá teorii, že v poslední době modernizace opravdu nabrala velmi rychlé tempo. Této oblasti se věnuje například sociolog Jan Keller (2007), jenž je v této práci také citován. Tato teorie se tedy potvrdila. Nejslabší závislostí naopak disponuje proměnná ptající se ochotu zvýšení pracovního tempa.

Tabulka 4: Výstup z programu IBM SPSS – Kanonická korelace v analýze č.1

	Function		
	1	2	3
Do jaké míry souhlasíte s výrokem - Naše společnost se mění tak rychle, že to člověk ani nestačí sledovat?	,797	,581	,164
Ohlédnete-li se zpět, myslíte si, že změna společenského systému u nás, k níž došlo v roce 1989, stála za to?	-,402	,284	,888
Vy sám/sama jste ochoten/na v případě potřeby zvýšit své pracovní tempo?	-,389	,716	-,606

Zdroj: Vlastní práce, ČSDA

Tabulka kanonické struktury neboli kanonických nákladů, představující korelaci mezi vysvětlujícími proměnnými a rozměry vytvořené diskriminační funkce, předchozí výsledek potvrzuje. Nejvyšší hodnoty jednotlivých proměnných jsou označeny hvězdičkou. Na druhou a třetí funkci není třeba brát zřetel, kvůli slabé diskriminační schopnosti.

Tabulka 5: Výstup z programu IBM SPSS – Kanonická struktura v analýze č.1

	Function		
	1	2	3
Do jaké míry souhlasíte s výrokem - Naše společnost se mění tak rychle, že to člověk ani nestačí sledovat?	,795*	,580	,174
Vy sám/sama jste ochoten/na v případě potřeby zvýšit své pracovní tempo?	-,462	,762*	-,453
Ohlédnete-li se zpět, myslíte si, že změna společenského systému u nás, k níž došlo v roce 1989, stála za to?	-,462	,413	,785*

Zdroj: Vlastní práce, ČSDA

Na základě výsledků diskriminační analýzy byla provedena následná klasifikace, která byla z 28,5 % úspěšná. V této tabulce by byl ideální výsledek, kdyby byly hodnoty pro jednotlivé proměnné pouze na diagonále.

Nejvíce se klasifikace vydařila u respondentů, kteří na vysvětlovanou proměnnou odpověděli „rozhodně souhlasím“ a „rozhodně nesouhlasím“ a to konkrétně z 55,1 % a 56,1 %. Naopak největší neúspěch zaznamenal program u nejpočetnější možnosti „spíše souhlasím“ a možná právě i proto jsou výsledky samotné diskriminační analýzy poněkud nepříznivé. Výsledkem tedy je, že u osob, kteří ve vysvětlované otázce odpověděli jednou z „krajních“ možností, se dá lépe zjistit určitá závislost na vysvětlujících proměnných pomocí které, je možné tyto osoby následně klasifikovat.

Tabulka 6: Výstup z programu IBM SPSS – Klasifikace do jednotlivých skupin v analýze č.1

		Do jaké míry souhlasíte s následujícími výroky - Tempo modernizace západních společností je nesmyslně vysoké?	Predicted Group Membership				Total
			Rozhodně souhlasím	Spíše souhlasím	Spíše nesouhlasím	Rozhodně nesouhlasím	
Original	Count	Rozhodně souhlasím	179	29	52	65	325
		Spíše souhlasím	293	76	151	202	722
		Spíše nesouhlasím	177	51	155	330	713
		Rozhodně nesouhlasím	71	21	53	185	330
		Ungrouped cases	134	21	25	83	263
%		Rozhodně souhlasím	55,1	8,9	16,0	20,0	100,0
		Spíše souhlasím	40,6	10,5	20,9	28,0	100,0
		Spíše nesouhlasím	24,8	7,2	21,7	46,3	100,0
		Rozhodně nesouhlasím	21,5	6,4	16,1	56,1	100,0
		Ungrouped cases	51,0	8,0	9,5	31,6	100,0

a. 28,5% of original grouped cases correctly classified.

Zdroj: Vlastní práce, ČSDA

Celkově se je tedy možné říci, že existuje silnější závislost mezi proměnnými, že tempo modernizace je velmi vysoké a tím, že se společnost mění opravdu rychle a člověk ji ani nestačí sledovat. Dalo by se také říci, že lidé, kteří mají na tempo modernizace jednoznačný názor se dají relativně dobře klasifikovat a využít v následných analýzách. Zároveň se podařilo potvrdit, že lidé už začali neustále se zrychlující tempo modernizace společnosti vnímat, což potvrzuje teorie řečené sociologem Janem Kellerem.

4.3 Analýza globalizace v rámci modernizace

V teoretické části bylo řečeno, že jestli nějaký jev úzce souvisí s modernizací společnosti, je to globalizace. Proto byla v následující analýze jako vysvětlovaná proměnná vybrána otázka „*Prosím, řekněte mi, zda jste pro modernizaci, i kdyby vedla k silné vlně přistěhovalectví z východu?*“. Naopak jako vysvětlující proměnné byly zvoleny otázky: „*Prosím, řekněte mi, zda jste pro modernizaci, i kdyby vedla ke stěhování lidí za prací?*“, „*Prosím, řekněte mi, zda jste pro modernizaci, i kdyby vedla k vytlačování některých skupin lidí na okraj společnosti?*“ a „*Vy sám/sama jste ochoten/na v případě potřeby odstěhovat se za prací do jiného města?*“.

Obecně by se zde dalo čekat, že lidem, kterým nevadí přistěhovalecká vlna z východu, nebude vadit ani obecně stěhování za prací.

Z analýzy bylo vyřazeno 127 respondentů, což je 5,4 % z celkového počtu respondentů. Jednalo se opět o odlehlé proměnné, které by zkreslovaly výsledek.

Tabulka 7: Výstup z programu IBM SPSS – Analýza proměnných v analýze č.2

Unweighted Cases		N	Percent
Valid		2226	94,6
Excluded	Missing or out-of-range group codes	127	5,4
	At least one missing discriminating variable	0	,0
	Both missing or out-of-range group codes and at least one missing discriminating variable	0	,0
	Total	127	5,4
Total		2353	100,0

Zdroj: Vlastní práce, ČSDA

Zajímavé je zde rozložení jednotlivých odpovědí uvedený v tabulce v příloze. Patrná je hlavně markantní „převaha“ možnosti „rozhodně ne“. Důvodem toho, že zde občané ČR nechtějí přistěhovalce z východu, mohou být například obavy z nárůstu kriminality, patriotismus či strach o ztrátu pracovního místa. Toto zkoumání by však vydalo na samostatnou analýzu v rámci přístupu k přistěhovalcům.

Z testování jednotlivých funkcí opět vyšla nejlépe první funkce. U dalších dvou je velikost diskriminační schopnosti opravdu velmi malá. Tuto hodnotnost potvrzuje i Wilksovo lamda rozdělení, kde hodnota menší než 0,05 ve sloupci Sig. značí, že je zde v podstatě minimální šance, aby zjištěná závislost mezi proměnnými byla náhodná.

Tabulka 8: Výstup z programu IBM SPSS – Testování diskriminační schopnosti v analýze č.2

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	,135 ^a	96,5	96,5	,345
2	,005 ^a	3,5	100,0	,070
3	,000 ^a	,0	100,0	,005

Zdroj: Vlastní práce, ČSDA

Tabulka 9: Výstup z programu IBM SPSS – Wilksovo lambda rozdělení v analýze č.2

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 3	,876	293,088	9	,000
2 through 3	,995	10,869	4	,028
3	1,000	,054	1	,816

Zdroj: Vlastní práce, ČSDA

Pozitivní korelace u prvních dvou vysvětlujících proměnných značí, že když respondenti odpovídají negativně na otázku migrace z východu, z poloviny odpovídají negativně i na to, zda by byli pro modernizaci i v případě, že by vedla ke stěhování za práci a zdali by byli pro vytlačování některých skupin na okraj společnosti. Koeficient u poslední otázky je velmi malý.

U těchto výsledků je tedy možné konstatovat, že občané české republiky jsou v tomto ohledu „solidární“ a nechtěli by, aby zvýšení modernizace mělo za následek odsunutí jednotlivých skupin obyvatel na okraj společnosti.

Tabulka 10: Výstup z programu IBM SPSS – Kanonická korelace v analýze č.2

	Function		
	1	2	3
Prosím, řekněte mi, zda jste pro modernizaci, i kdyby vedla ke stěhování lidí za prací?	,583	-,073	-,852
Prosím, řekněte mi, zda jste pro modernizaci, i kdyby vedla k vytlačování některých skupin lidí na okraj společnosti?	,626	-,493	,645
Vy sám/sama jste ochoten/na v případě potřeby odstěhovat se za prací do jiného města?	,212	,965	,306

Zdroj: Vlastní práce, ČSDA

Tabulka kanonické struktury tyto výsledky opět potvrzuje. V první funkci jsou obsaženy největší koeficienty pro první dvě vysvětlující proměnné.

Tabulka 11: Výstup z programu IBM SPSS – Kanonická struktura v analýze č.2

	Function		
	1	2	3
Prosím, řekněte mi, zda jste pro modernizaci, i kdyby vedla k vytlačování některých skupin lidí na okraj společnosti?	,762*	-,341	,550
Prosím, řekněte mi, zda jste pro modernizaci, i kdyby vedla ke stěhování lidí za prací?	,736*	,053	-,674
Vy sám/sama jste ochoten/na v případě potřeby odstěhovat se za prací do jiného města?	,445	,866*	,230

Zdroj: Vlastní práce, ČSDA

V následné klasifikaci se potvrdilo, že i když byla nejvíc vypovídající pouze první funkce, stačilo to k tomu, aby byla klasifikace úspěšná v 42,5 %. Ze 70,3 % se to povedlo u první možnosti, která však byla početně daleko slabší než poslední možnost. Poslední možnost však vyšla také přijatelně a z 57,8 % se proměnné původně spadající do této možnosti povedlo správně zařadit na základě výsledků z diskriminační analýzy.

Nejmenší shoda opět panovala ve středních možnostech, což potvrzuje, že respondenti, kteří nemají jasný názor ve vysvětlované proměnné, mohou mít jakýkoliv názor ve vysvětlujících otázkách.

Tabulka 12: Výstup z programu IBM SPSS – Klasifikace do jednotlivých skupin v analýze č.2

		Prosím, řekněte mi, zda jste pro modernizaci, i kdyby vedla k silné vlně přistěhovalectví z východu?	Predicted Group Membership				Total
			rozhodně ano	spíše ano	spíše ne	rozhodně ne	
Original	Count	rozhodně ano	83	20	3	12	118
		spíše ano	124	72	37	56	289
		spíše ne	125	191	142	239	697
		rozhodně ne	116	193	164	649	1122
		Ungrouped cases	13	8	10	96	127
%		rozhodně ano	70,3	16,9	2,5	10,2	100,0
		spíše ano	42,9	24,9	12,8	19,4	100,0
		spíše ne	17,9	27,4	20,4	34,3	100,0
		rozhodně ne	10,3	17,2	14,6	57,8	100,0
		Ungrouped cases	10,2	6,3	7,9	75,6	100,0

a. 42,5% of original grouped cases correctly classified.

Zdroj: Vlastní práce, ČSDA

Tato analýza ukázala jistou míru patriotismu a sociálního citění s ostatními, když občané České republiky nesouhlasí s modernizací, kdyby vedla k odstrčení určitých skupin na okraj společnosti a zároveň se nechtějí v rámci modernizace příliš stěhovat za prací. Tyto dvě otázky jsou navíc provázány s možností, zda by respondenti byli pro modernizaci, která by zapříčinila migraci lidí z blízkého východu.

Výsledkem tedy je, že lidé, kteří nejsou pro migraci lidí z východu, by zároveň v rámci modernizace neobětovali určité skupiny obyvatel, které by modernizace odsunula na okraj společnosti.

Lidé v rámci modernizace neztrácejí empatii se sociálně slabšími skupinami, a naopak nechtějí obětovat svůj život a svoji práci a stěhovat za prací kamkoliv jinam, aby stejnou práci místo nich vykonávala levná pracovní síla z východu. Toto se vymyká teorii napsané v kapitole 3.3.1 a to sice té, že lidé se v první vlně globalizace stěhovali z menších měst do větších a tím velmi prohlubovali bídu na venkově, a tak i vytlačovali určité skupiny obyvatel na okraj společnosti.

4.4 Analýza postojů k modernizaci

V této analýze bude snahou zjistit jaký typ člověka je pro modernizaci a jaký naopak pro život založený na tradicích? Proto byla jako vysvětlovaná proměnná v následné analýze zvolena otázka: „*Jakými zásadami by se měla řídit společnost, v níž byste chtěl/a žít - 1. varianta Upřednostňovat dodržování tradic? - 2. varianta Upřednostňovat modernizační změny?*“.

Jako první základní vysvětlující proměnná bylo zvoleno „*Pohlaví*“ respondenta. Dále je zde brán v potaz předpoklad, že o modernizaci se daleko více zajímají lidé s vysokoškolským vzděláním, a proto byla jako druhá vysvětlující proměnná zvolena otázka: „*Jaké je vaše nejvyšší ukončené školní vzdělání?*“. A jako poslední proměnná byl zvolen „*Rodinný stav*“.

Z pozorování bylo vyřazeno 93 vzorků, aby byl odstraněn vliv odlehlých hodnot na výsledky testu.

Tabulka 13: Výstup z programu IBM SPSS – Analýza proměnných v analýze č.3

Unweighted Cases		N	Percent
Valid		2260	96,0
Excluded	Missing or out-of-range group codes	93	4,0
	At least one missing discriminating variable	0	,0
	Both missing or out-of-range group codes and at least one missing discriminating variable	0	,0
	Total	93	4,0
Total		2353	100,0

Zdroj: Vlastní práce, ČSDA

Velmi vyrovnané je rozložení odpovědí k jednotlivým možnostem, uvedených v tabulce v příloze. Největší zastoupení mají opět nejednoznačné možnosti, a proto se dá, na základě předchozích analýz, předpokládat velmi špatně předpověditelná závislost a také následná klasifikace do jednotlivých skupin. Nicméně o trochu větší zastoupení mají možnosti vyjadřující se spíše pro modernizaci než pro tradici.

V tabulce znázorňující diskriminační schopnost jednotlivých funkcí podle předpokladů vyšly slabší hodnoty u hodnot značících právě tuto schopnost. Jedinou

drobnou výjimkou je první funkce. Tato funkce má nejvyšší vypovídající schopnost ze všech tří funkcí přesně tak, jak značí druhý sloupec v tabulce č. 14.

Tabulka 14: Výstup z programu IBM SPSS – Testování diskriminační schopnosti v analýze č.3

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	,033 ^a	93,0	93,0	,179
2	,002 ^a	6,1	99,0	,046
3	,000 ^a	1,0	100,0	,019

Zdroj: Vlastní práce, ČSDA

Slabší diskriminační schopnost jednotlivých funkcí potvrdila i tabulka Wilksova lambda rozdělení. Většina hodnot se velmi podstatě blíží 1 a testování třetí funkce se dokonce přímo rovná 1. Tyto hodnoty jsou pro analýzu ne příliš dobré. Co opět vyznívá pozitivně pro tuto analýzu, je velmi nízká hodnota náhodnosti vyhledané závislosti.

Tabulka 15: Výstup z programu IBM SPSS – Wilksovo lambda rozdělení v analýze č.3

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 3	,965	79,363	9	,000
2 through 3	,997	5,659	4	,226
3	1,000	,782	1	,377

Zdroj: Vlastní práce, ČSDA

V tabulce jednotlivých koeficientů je největší vliv přikládán rodinnému stavu. Negativní korelace u pohlaví říká, že muži jsou spíše pro modernizaci a ženy spíše pro dodržování tradic. Těmto hodnotám se však nedá přikládat příliš velká důležitost s přihlédnutím na výsledky celkového modelu. Tabulka kanonických nákladů byla vzhledem k ne příliš dobrým výsledkům celého modelu v tomto případě vynechána.

Tabulka 16: Výstup z programu IBM SPSS – Kanonická korelace v analýze č.3

	Function		
	1	2	3
Jaké je vaše nejvyšší ukončené školní vzdělání?	,433	,819	,382
Pohlaví	-,521	-,174	,837
Rodinný stav	,766	-,520	,381

Zdroj: Vlastní práce, ČSDA

Na základě výsledků diskriminační analýzy se v tom případě nepodařilo kvalitně zařadit proměnné do jejich správných kategorií a celkově byl zaznamenán úspěch jen v 26,6 % případů. Nejvíce se to povedlo u možností „rozhodně první“, která značila respondenty, kteří dávají přednost spíše tradicím na úkor modernizace. Tyto výsledky jsou pochopitelné vzhledem k celkově špatné analýze, která nenašla jakoukoliv silnější závislost mezi jednotlivým proměnnými.

Tabulka 17: Výstup z programu IBM SPSS – Klasifikace do jednotlivých skupin v analýze č.3

		Jakými zásadami by se měla řídit společnost, v níž byste chtěli/a žít-1. varianta Upřednostňovat dodržování tradic?-2. varianta Upřednostňovat modernizační změny?	Predicted Group Membership				Total
			Rozhodně první	Spíše první	Spíše druhá	Rozhodně druhá	
Original	Count	Rozhodně první	148	43	46	67	304
		Spíše první	329	126	122	189	766
		Spíše druhá	244	119	137	228	728
		Rozhodně druhá	115	64	93	190	462
		Ungrouped cases	24	15	14	40	93
%		Rozhodně první	48,7	14,1	15,1	22,0	100,0
		Spíše první	43,0	16,4	15,9	24,7	100,0
		Spíše druhá	33,5	16,3	18,8	31,3	100,0
		Rozhodně druhá	24,9	13,9	20,1	41,1	100,0
		Ungrouped cases	25,8	16,1	15,1	43,0	100,0

a. 26,6% of original grouped cases correctly classified.

Zdroj: Vlastní práce, ČSDA

Jak již bylo řečeno výše, klasifikace na základě výsledků diskriminační analýzy nedopadla příznivě. Není možné jednoznačně říci, kteří lidé vyznávají názor toho, že by se ČR měla zaměřit na modernizaci než na kulturní tradice. Tento poznatek je i přes nepřítel celkové analýzy hodnotný a říká, že není možné posoudit, zdali vzdělanější lidé podporují

modernizaci více než lidé s menším vzděláním. Stejně tak není možné s větší jistotou říci, jestli o modernizaci jako takovou stojí spíše muži než ženy.

4.5 Analýza postojů k ekonomickému růstu

V teoretické části této práce je uvedeno, že modernizace s sebou nese i riziko devastace venkova a životního prostředí. Co naopak může být příznivým dopadem modernizace, je ekonomický růst. Proto budou nyní tyto dvě možnosti postaveny proti sobě a tato analýza se pokusí odhalit, jestli se dá říci, které typy osob by volili spíše cestu ekonomického růstu a kteří naopak péči o životní prostředí.

Jako vysvětlovaná proměnná byla zvolená otázka „*Jakými zásadami by se měla řídit společnost, v níž byste chtěl/a žít - 1. varianta Starat se hlavně o dobré životní prostředí? - 2. varianta Starat se hlavně o ekonomický růst?*“. Jako vysvětlující bylo opět vybráno „*Pohlaví*“ a dále pak otázky „*Jakými zásadami by se měla řídit společnost, v níž byste chtěl/a žít - 1. varianta Upřednostňovat dodržování tradic? - 2. varianta Upřednostňovat modernizační změny?*“ a „*Které variantě dáváte přednost - 1. varianta Rozhodovat spíše podle názorů veřejnosti? - 2. varianta Rozhodovat spíše podle odborníků?*“.

Z analýzy bylo vyřazeno 91 proměnných ze stejného důvodu jako v předchozích případech, a to sice aby byly odstranili odlehlé hodnoty, které by mohly ovlivňovat normalitu rozdělení.

Tabulka 18: Výstup z programu IBM SPSS – Analýza proměnných v analýze č.4

Unweighted Cases		N	Percent
Valid		2262	96,1
Excluded	Missing or out-of-range group codes	91	3,9
	At least one missing discriminating variable	0	,0
	Both missing or out-of-range group codes and at least one missing discriminating variable	0	,0
	Total	91	3,9
Total		2353	100,0

Zdroj: Vlastní práce, ČSDA

Na druhé tabulce, vztahující se k této analýze, která je uvedena v příloze, je vidět zjištění, že lidé se spíše klaní k péči o životní prostředí než o ekonomický růst. Největší

počet opět zaujímají nejednoznačné možnosti, ale to se u takto složitých typů otázek dalo předpokládat. Jednoznačné odpovědi vyplnilo na jedné straně 409 a na druhé straně 317 respondentů.

Na tabulce týkající se závislosti značící diskriminační schopnost dopadla nejlépe první funkce, která má i velmi slušnou hodnotu kanonické korelace, jež by mohla značit určitou „správnost“ vyhledané závislosti. Samozřejmě opět platí, že čím více se číslo přibližuje v 1, tím lepší je to pro konečnou formulaci závěrů na základě výzkumu.

Tabulka 19: Výstup z programu IBM SPSS – Testování diskriminační schopnosti v analýze č.4

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	,064 ^a	96,8	96,8	,246
2	,001 ^a	2,2	99,0	,038
3	,001 ^a	1,0	100,0	,026

Zdroj: Vlastní práce, ČSDA

Tabulka Wilksova lambda rozdělení potvrzuje předešlé tvrzení o síle závislosti, především pak sloupce týkající se přímo Wilksova lambda rozdělení a také p-hodnota v posledním sloupci značící, že se nejedná o náhodnou závislost.

Tabulka 20: Výstup z programu IBM SPSS – Wilksovo lambda rozdělení v analýze č.4

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 3	,938	145,416	9	,000
2 through 3	,998	4,837	4	,304
3	,999	1,500	1	,221

Zdroj: Vlastní práce, ČSDA

Tabulka diskriminačních koeficientů potvrzuje předpoklad, že pro ekonomický růst se vyslovily osoby, které zároveň upřednostňují modernizaci, a naopak lidé lpící na tradicích se vyjádřili pro šetrné životní prostředí. Tento výsledek byl očekávaný. Za neočekávaný by se však dala označit velmi slabá závislost u dalších dvou otázek. Opět, stejně jako v předchozí analýze, se ukázalo, že není možné určit podle pohlaví osoby, které jsou pro modernizaci a růst ekonomiky a které naopak pro zachování tradic a životního prostředí. Tento výsledek ukazuje, že ani muži ani ženy na tuto problematiku nemají jasný názor.

Tabulka 21: Výstup z programu IBM SPSS – Kanonická korelace v analýze č.4

	Function		
	1	2	3
Pohlaví	-,277	,520	,809
Jakými zásadami by se měla řídit společnost, v níž byste chtěl/a žít-1. varianta Upřednostňovat dodržování tradic?-2. varianta Upřednostňovat modernizační změny?	,892	-,150	,445
Které variantě dáváte přednost- 1.varianta Rozhodovat spíše podle názorů veřejnosti?-2. varianta Rozhodovat spíše podle odborníků?	,236	,863	-,465

Zdroj: Vlastní práce, ČSDA

Výsledky následné klasifikace pro tuto analýzu jsou rozporuplné. Velmi dobrá je zařaditelnost krajních hodnot, především pak těch respondentů, kteří se vyjádřili pro životní prostředí, ale naopak velmi slabá je klasifikace středových hodnot, obzvláště pak osob, které zvolili možnost „spíše druhá“ varianta.

Tabulka 22: Výstup z programu IBM SPSS – Klasifikace do jednotlivých skupin v analýze č.4

		Jakými zásadami by se měla řídit společnost, v níž byste chtěl/a žít-1. varianta Starat se hlavně o dobré životní prostředí?-2.varianta Starat se hlavně o ekonomický růst?	Predicted Group Membership				Total
			Rozhodně první	Spíše první	Spíše druhá	Rozhodně druhá	
Original	Count	Rozhodně první	223	78	42	66	409
		Spíše první	359	220	80	208	867
		Spíše druhá	142	165	111	251	669
		Rozhodně druhá	68	44	68	137	317
		Ungrouped cases	20	14	15	42	91
	%	Rozhodně první	54,5	19,1	10,3	16,1	100,0
		Spíše první	41,4	25,4	9,2	24,0	100,0
		Spíše druhá	21,2	24,7	16,6	37,5	100,0
		Rozhodně druhá	21,5	13,9	21,5	43,2	100,0
		Ungrouped cases	22,0	15,4	16,5	46,2	100,0

a. 30,5% of original grouped cases correctly classified.

Zdroj: Vlastní práce, ČSDA

Výsledek této analýzy ukázal na dva fakty. První z nich je ten, že lidé podporující modernizaci jsou zároveň i lidé podporující ekonomický růst na úkor životního prostředí a za druhé se zároveň opět ukázalo, respektive potvrdilo, že není možné osoby klasifikovat podle jednoduchých dotazníkových otázek, jako je například pohlaví respondenta.

4.6 Analýza závislosti přeměny společnosti na tempu modernizace – Metoda CART

V následujících kapitolách budou provedeny stejné analýzy jako v předchozích případech s tím rozdílem, že nyní bude využita zmíněná metoda klasifikačních a regresních stromů, zkráceně CART.

U této metody není potřeba řešit odlehlé hodnoty, jelikož metoda pracuje na bázi Giniho indexu, ale v rámci toho, aby strom v analýze nebyl příliš „košatý“, byly tyto hodnoty opět odstraněny.

Podmínka týkající se dělení počtu pozorování v kořenu stromu, v angličtině tzv. Parent node, zůstala zachována na přednastavených 100, stejně tak podmínka Child node, značící podmínku pro počty pozorování ve větvích stromu. Tato hodnota byla přednastavena na 50.

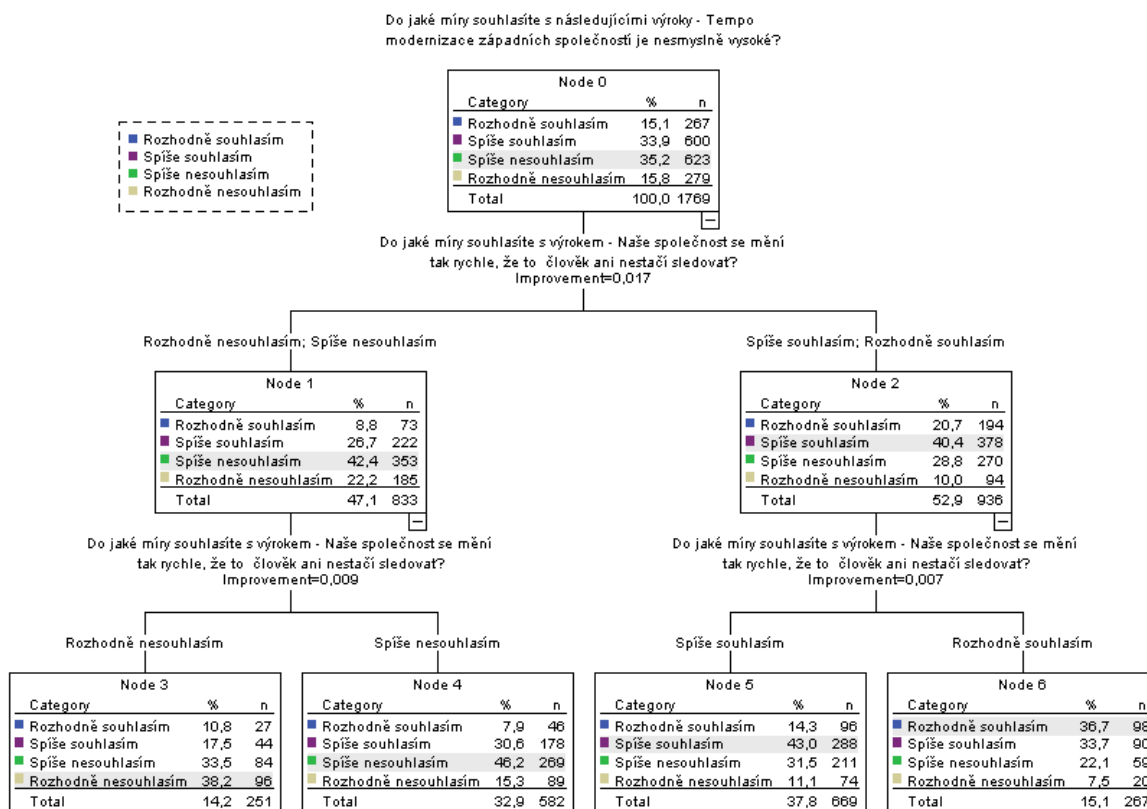
Výsledkem analýzy je 5 úroňový klasifikační strom. Pro jeho velikost bude do této práce vkládán postupně.

Na vrcholu celého stromu je vysvětlovaná proměnná a rozdělení do jednotlivých možností tak, jak na ni respondenti odpověděli.

V první úrovni už se pak strom dělí podle první vysvětlující proměnné. Větve stromu jsou rozděleny na dva extrémy, v tomto případě na kladné a negativní odpovědi. Ve druhé úrovni jsou pak větve opět rozděleny do dalších dvou extrémních protikladů z předcházející odpovědi přesně tak, jak je napsáno v metodické části práce.

Tato úroveň je pro popis podstatná, jelikož zde končí první vysvětlující proměnná. V této úrovni jsou vidět předpokládané výsledky v tom, že lidé, kteří na vysvětlovanou otázku odpověděli jednou ze čtyř názorových možností, pak zvolili i stejnou možnost v první vysvětlující otázce. Jinými slovy lidé, kteří rozhodně nesouhlasí s tvrzením, že tempo modernizace západní společnosti je příliš vysoké, pak také rozhodně nesouhlasí s tvrzením, že společnost se mění příliš rychle.

Obrázek 5: Výstup z programu IBM SPSS – První dvě úrovně v CART – analýza č.1



Zdroj: Vlastní práce, ČSDA

Na obrázku č. 15 uvedeném v příloze je vidět, že v levé a pravé větvi se strom dělí podle proměnné týkající se zvýšení pracovního tempa, ale naopak v prostřední větvi metoda CART vyhodnotila jako nejlepší větvit strom podle proměnné ptající se na názor na politickou změnu v roce 1989.

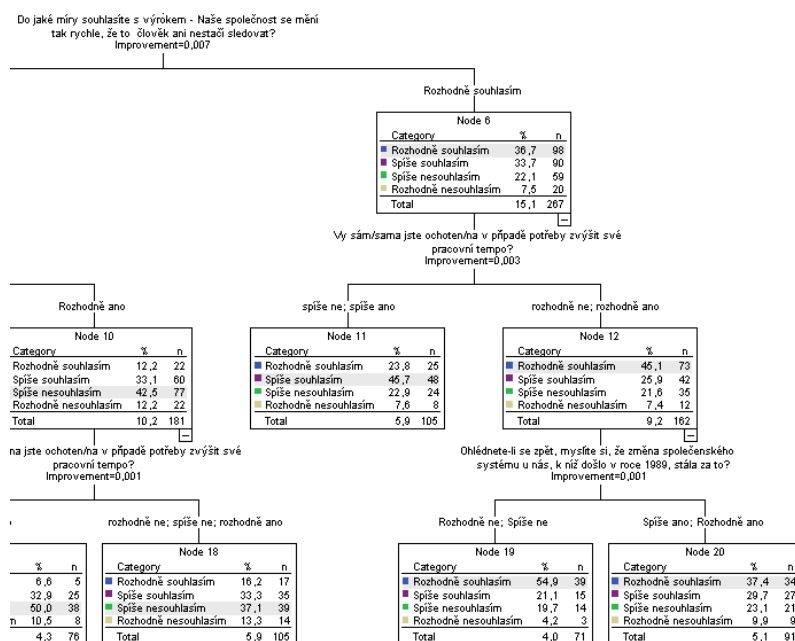
Potvrzuje se, že lidé, kteří by nebyli ochotni zvýšit své pracovní tempo, spíše souhlasí s tím, že tempo modernizace západu je nesmyslně vysoké, tak jak bylo již vidno z diskriminační analýzy. Naopak možnost „rozhodně ano“ skončila hned ve třetí větvi a dále už se „nerozpadala“ do dalších úrovní. Tito respondenti si podle předpokladu nemyslí, že tempo modernizace západních společností je příliš vysoké.

Zatímco levá větev dále dělí 372 respondentů podle další vysvětlující proměnné, prostřední větev zvolila odlišný postup a nejprve dělila podle názoru na rok 1989 a až poté

podle otázky na zvýšení pracovního tempa. Je tedy dobré se věnovat prostřední větvi, která pracuje s větším počtem respondentů, konkrétně s 488. V poslední úrovni této větve je vidět, že lidé, kteří rozhodně souhlasí s tím, že změna v roce 1989 stála za to, nesouhlasí s tvrzením, že tempo modernizace západu je příliš vysoké.

Na pravé větvi lze vidět, že metoda CART již ani nevětvila strom do finální 5. úrovně z důvodu kritéria pro počet pozorování ve větvích. Na pravé větvi, která vycházela z krajních odpovědí na otázku pracovního tempa, je vidět, že ze 162 respondentů jich 91 zvolil kladnou odpověď na otázku politické změny roku 1989 a tito lidé pak i souhlasili s tvrzením, že tempo modernizace je příliš vysoké.

Obrázek 6: Výstup z programu IBM SPSS – Pravá větev – analýza č.1



Zdroj: Vlastní práce, ČSDA

Nyní stejně jako v diskriminační analýze, i zde bude provedena na základně předešlého rozhodovacího stromu klasifikace proměnných do jednotlivých kategorií. Tabulka Risk popisuje, z kolika procent se klasifikace nezdařila. Klasifikace se nezdařila z 54,7 %, což potvrzuje i následná klasifikační tabulka, jejíž poslední číslo 45,3 % je počet správně zařazených proměnných.

Tabulka 23: Výstup z programu IBM SPSS – Tabulka Risk – analýza č.1

Estimate	Std. Error
,547	,012

Growing Method:
 CRT
 Dependent Variable:
 Do jaké míry
 souhlasíte s
 následujícími výroky -
 Tempo modernizace
 západních
 společností je
 nesmyslně vysoké?

Zdroj: Vlastní práce, ČSDA

Tabulka 24: Výstup z programu IBM SPSS – Klasifikace podle CART – analýza č.1

Observed	Predicted				Percent Correct
	Rozhodně souhlasím	Spíše souhlasím	Spíše nesouhlasím	Rozhodně nesouhlasím	
Rozhodně souhlasím	73	111	56	27	27,3%
Spíše souhlasím	42	306	208	44	51,0%
Spíše nesouhlasím	35	178	326	84	52,3%
Rozhodně nesouhlasím	12	70	101	96	34,4%
Overall Percentage	9,2%	37,6%	39,1%	14,2%	45,3%

Zdroj: Vlastní práce, ČSDA

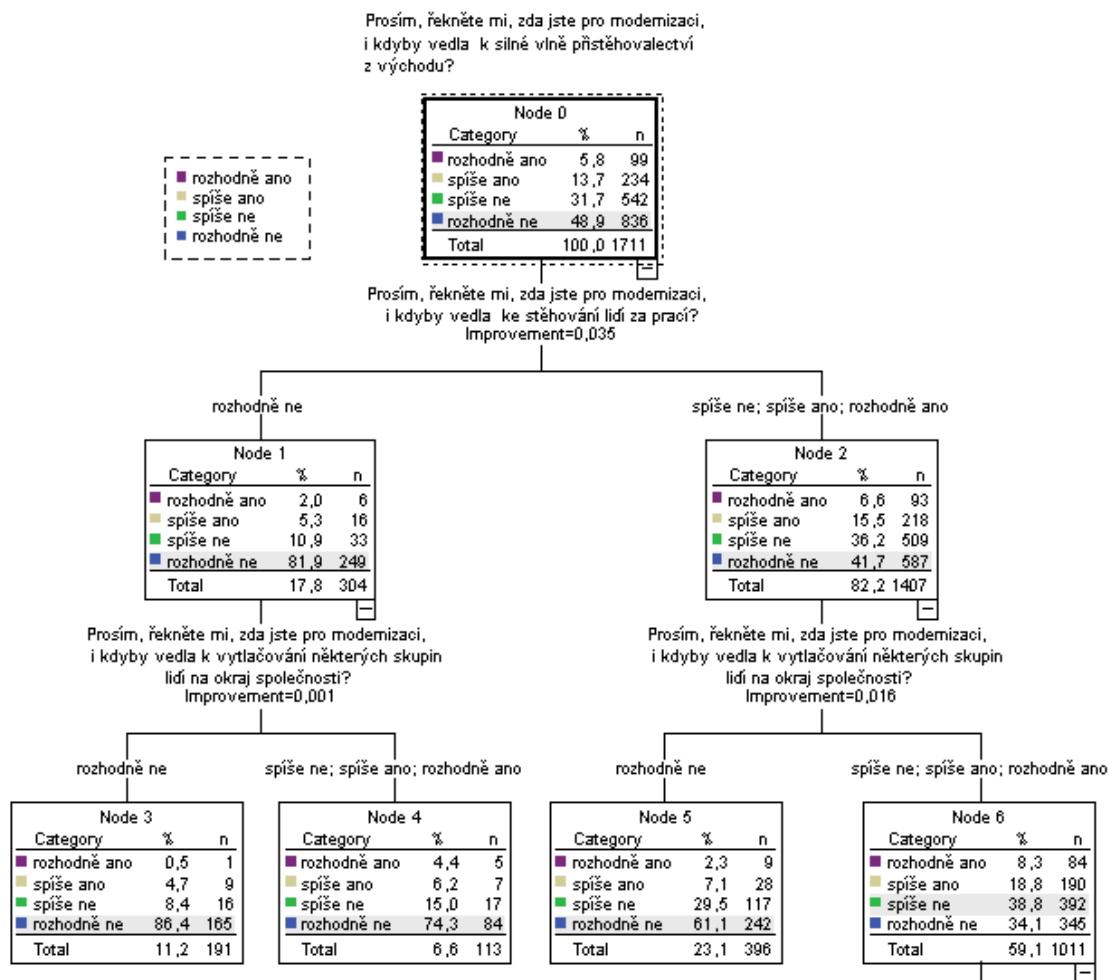
Následná klasifikace vyšla lépe než v případě diskriminační analýzy. Klasifikační strom v tomto případě dokázal lépe odhadnout zařazení jednotlivých proměnných a jeho model dokázal tyto proměnné lépe klasifikovat.

4.7 Analýza globalizace v rámci modernizace – metoda CART

Při provedení další analýzy v rámci metody CART bylo opět zachováno původní nastavení Parent node a child node. Byly také odstraněny odpovědi „nevím“ a „neodpověděl/a“, nikoli však z důvodu odlehlých hodnot, ale z důvodu, aby výsledný klasifikační strom nebyl příliš košatý.

Na první úrovni stromu byla jako rozřazující proměnná zvolena otázka týkající se stěhování lidí za prací. Na první úrovni se také potvrdil výsledek z diskriminační analýzy, který říkal, že lidé, kteří se nechtějí stěhovat za prací, tak také zároveň nechtějí modernizaci, kdyby vedla k přistěhovalecké vlně z východu.

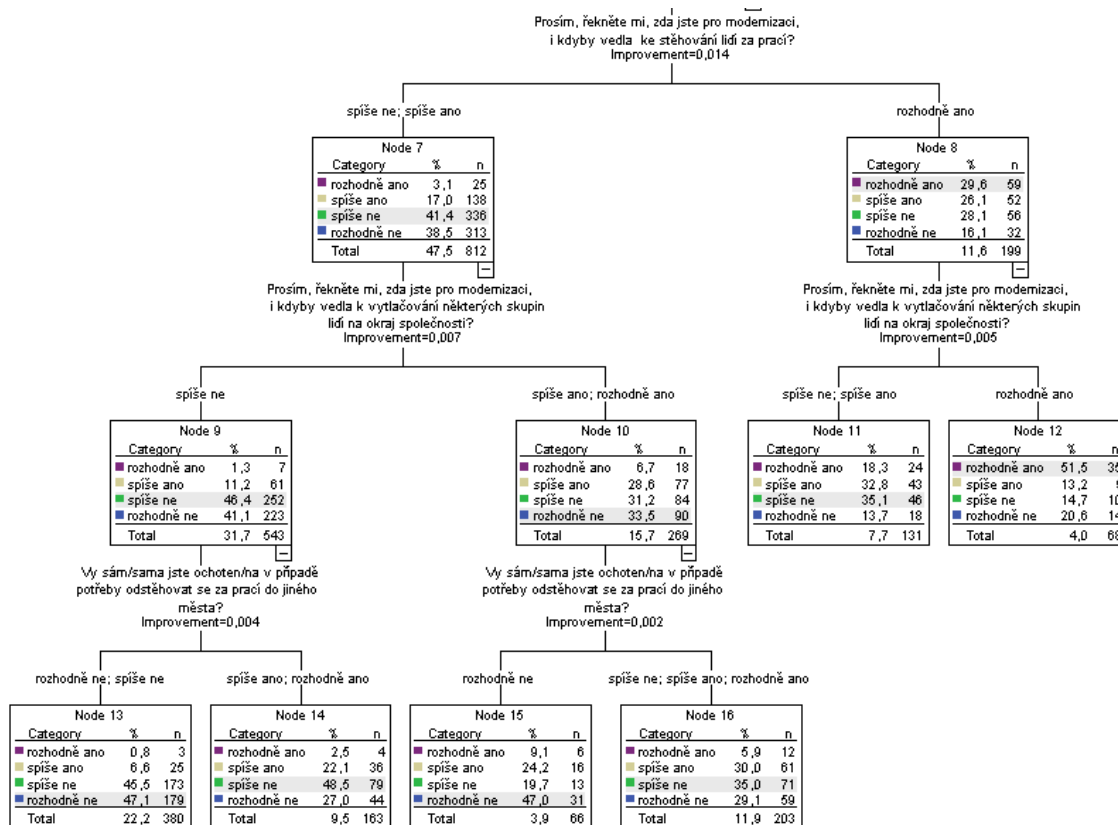
Obrázek 7: Výstup z programu IBM SPSS – První dvě úrovně CART – analýza č.2



Zdroj: Vlastní práce, ČSDA

Na druhé úrovni pak již metoda CART zvolila další vysvětlující proměnnou, a to sice otázku na vytlačování některých skupin na okraj společnosti právě v rámci modernizace. Zde se opět potvrdil výsledek z diskriminační analýzy, kdy lidé, kteří nechtějí vytlačovat některé skupiny obyvatel, zároveň nechtějí levnou pracovní sílu z východu z důvodu modernizace.

Obrázek 8: Výstup z programu IBM SPSS – Další úrovně v metodě CART – analýza č.2



Zdroj: Vlastní práce, ČSDA

Jelikož po druhé větvi nebylo splněno kritérium v rámci Parent node a child node, byla metodou zvolena další vysvětlující proměnná. Zajímavé však je, že ve čtvrté úrovni se opět vrátila proměnná týkající se vytlačování menšin.

V poslední úrovni už se objevuje i poslední vysvětlující proměnná a zde stojí za povšimnutí, že téměř v jakékoliv odpovědi na tuto otázku je i otázka na vysvětlovanou proměnnou negativní.

Tabulka 25: Výstup z programu IBM SPSS – Tabulka Risk – analýza č.2

Estimate	Std. Error
,455	,012

Growing Method:
 CRT
 Dependent Variable:
 Prosím, řekněte mi,
 zda jste pro
 modernizaci, i kdyby
 vedla k silné vlně
 přistěhovalectví
 z východu?

Zdroj: Vlastní práce, ČSDA

Na základě výsledků metody CART se podařilo správně zařadit proměnné v 54,5 % případů. Neúspěch byl pak zaznamenán v 45,5 % případů.

Tuto analýzu lze označit jako úspěšnou, jelikož se podařilo jednak potvrdit předpoklady řečené před samotnou analýzou, vycházející z teorie o modernizaci a dále pak také i výsledky z předchozí diskriminační analýzy. Klasifikace u metody CART skončila velmi dobře a proměnné se podařilo správně zařadit ve více jak polovině případů.

Tabulka 26: Výstup z programu IBM SPSS – Klasifikace podle CART – analýza č.2

Observed	Predicted				Percent Correct
	rozhodně ano	spíše ano	spíše ne	rozhodně ne	
rozhodně ano	35	0	40	24	35,4%
spíše ano	9	0	140	85	0,0%
spíše ne	10	0	196	336	36,2%
rozhodně ne	14	0	121	701	83,9%
Overall Percentage	4,0%	0,0%	29,0%	67,0%	54,5%

Zdroj: Vlastní práce, ČSDA

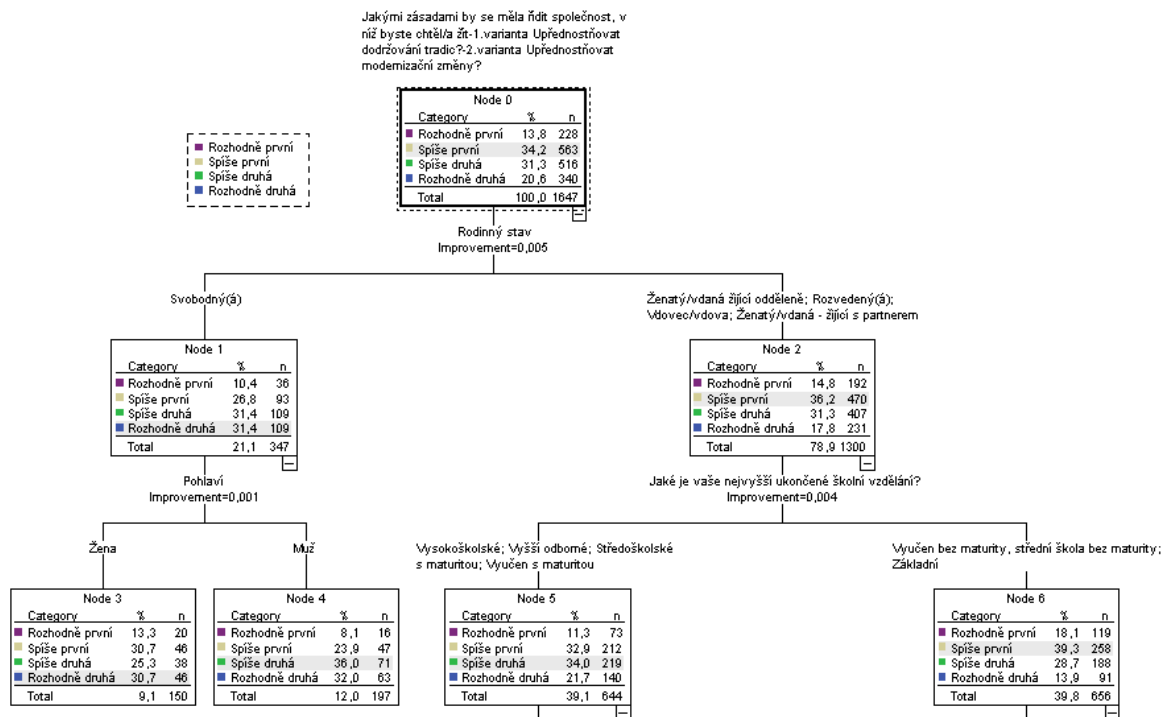
4.8 Analýza postojů k modernizaci – metoda CART

Analýza postojů k modernizaci, která v případě diskriminační analýzy vyšla nejhůře ze všech čtyř analýz, jež byly provedeny, bude nyní provedena pomocí metody CART.

Metoda v první větvi rozdělila odpovědi podle rodinného stavu respondentů. Na jednu stranu větve dala svobodné respondenty a na druhou stranu ostatní rodinné stavy. Ve druhé úrovni se pak levé stranu větvi podle pohlaví a pravá naopak podle nejvyššího ukončeného

školního vzdělání. Vysokoškolsky vzdělaní lidé jsou spíše pro modernizaci, ale rozdíly mezi odpověďmi v této kategorii jsou opravdu velmi malé.

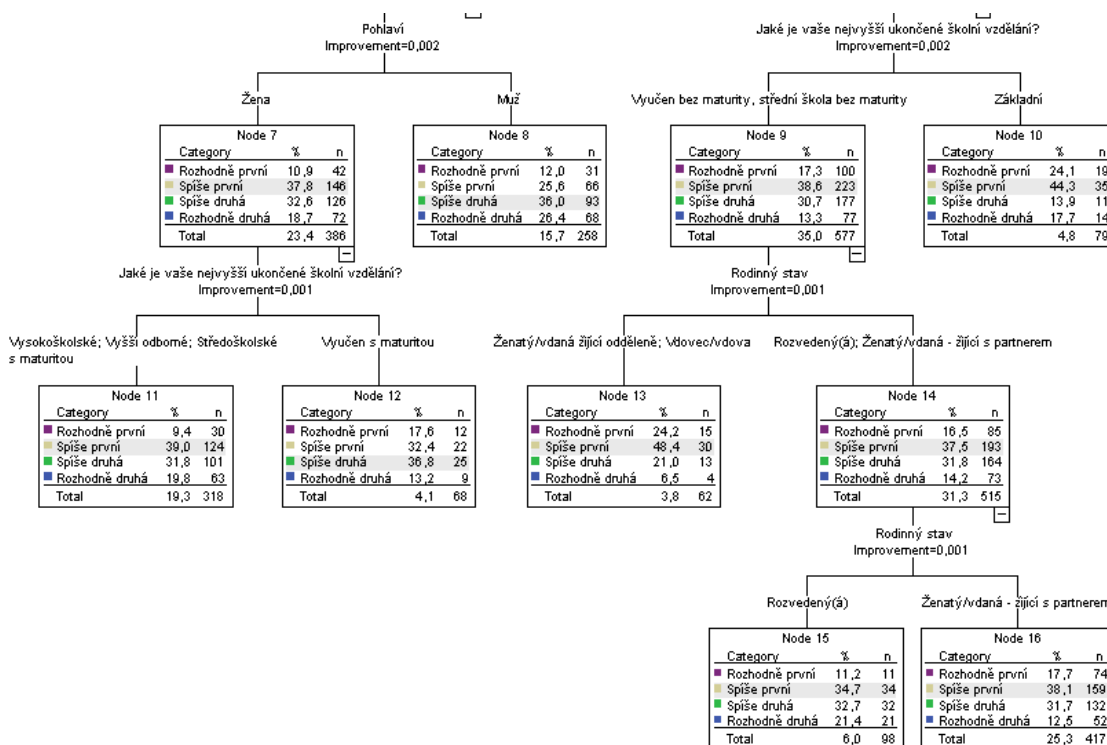
Obrázek 9: Výstup z programu IBM SPSS – První dvě úrovně metody CART – analýza č.3



Zdroj: Vlastní práce, ČSDA

Ve třetí úrovni se pak střední větev dělí podle pohlaví a pravá větev dále podle vzdělání. Vzdělání bez maturity se pak již dělí podle rodinného vztahu odpovídajícího. Jedná se o 577 respondentů. Lidé bez maturity, ať už jsou v jakémkoliv rodinném stavu, by raději volili dodržování nastolených tradic než upřednostňování modernizační změny.

Obrázek 10: Výstup z programu IBM SPSS – Další úrovně metody CART – analýza č.3



Zdroj: Vlastní práce, ČSDA

V následné klasifikační tabulce je vidět stejně nepříznivý výsledek, jako byl pro stejnou analýzu i v diskriminační analýze. Špatné zařazení proměnných se stalo v 62,5 % případech, a naopak správné zařazení se povedlo pouze u 37,5 % dotázaných.

Tabulka 27: Výstup z programu IBM SPSS Tabulka Risk – analýza č.3

Estimate	Std. Error
,625	,012

Growing Method:
 CRT
 Dependent Variable:
 Jakými zásadami by se měla řídit společnost, v níž byste chtěl/a žít-1.
 varianta
 Upřednostňovat dodržování tradic?-2.
 varianta
 Upřednostňovat modernizační změny?

Zdroj: Vlastní práce, ČSDA

Tabulka 28: Výstup z programu IBM SPSS – Klasifikace podle CART – analýza č.3

Observed	Predicted				Percent Correct
	Rozhodně první	Spiše první	Spiše druhá	Rozhodně druhá	
Rozhodně první	0	149	59	20	0,0%
Spiše první	0	382	135	46	67,9%
Spiše druhá	0	289	189	38	36,6%
Rozhodně druhá	0	154	140	46	13,5%
Overall Percentage	0,0%	59,1%	31,8%	9,1%	37,5%

Zdroj: Vlastní práce, ČSDA

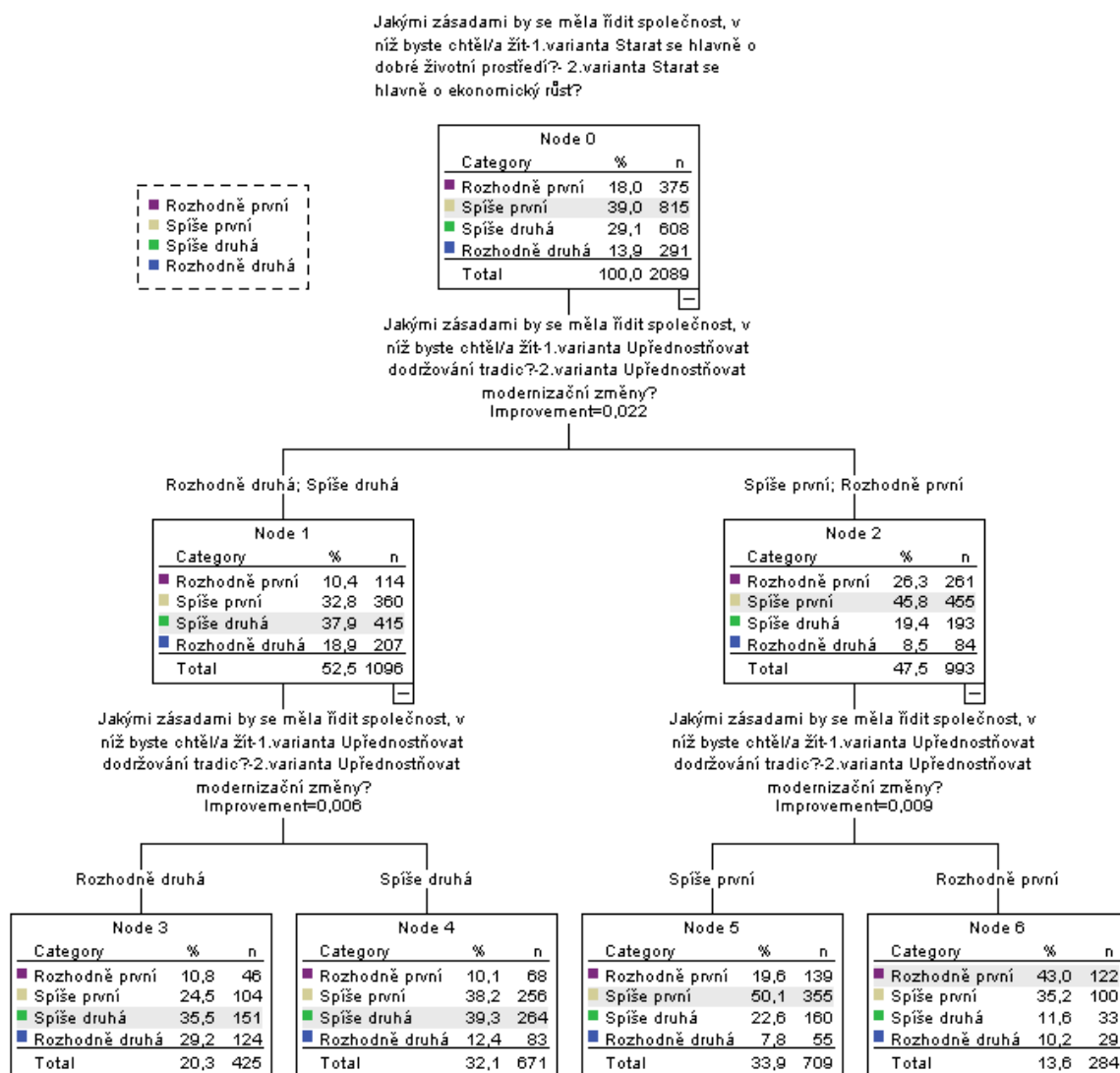
V rámci analýzy se podařilo potvrdit stejný výsledek jako v případě diskriminační analýzy, a to sice ten, že se dá velmi těžko určit, který občan je pro modernizaci a který naopak je jejím odpůrcem. Tento vztah k modernizaci není možné určit pomocí základních klasifikačních znaků, jako jsou pohlaví, rodinný vztah či vzdělání konkrétní osoby. Tento fakt potvrzuje, jak složitá oblast modernizace je a jak těžké je definovat její příznivce či odpůrce, což potvrzuje i slabý výsledek následné klasifikace.

4.9 Analýza postojů k ekonomickému růstu – metoda CART

Jako poslední je na řadě analýza postojů k ekonomickému růstu. Určitou závislost se sice diskriminační analýze podařilo najít, nicméně následná klasifikace se dostala pouze na 30,5 %. V tomto případě u metody CART však již musela být zvýšena hodnota Parent node na 150 a child node naopak snížena 30. Tato opatření byla uskutečněna z toho důvodu, aby bylo zabráněno přílišné košatosti stromu.

V prvních dvou úrovních se obě větve dělí podle jedné a té samé proměnné zaměřující se na rozdíl mezi dodržování tradic a modernizační proměnnou. Potvrzuje se výsledek jak z diskriminační analýzy, tak i z teoretické části této práce, kdy lidé, co stojí za modernizací, by se zároveň raději starali o ekonomický růst země. Protipól ekonomickému růstu je ještě více jednoznačnější a lidé trávající na tradicích by se raději starali o životní prostředí.

Obrázek 11: Výstup z programu IBM SPSS – První dvě úrovně metody CART – analýza č.4

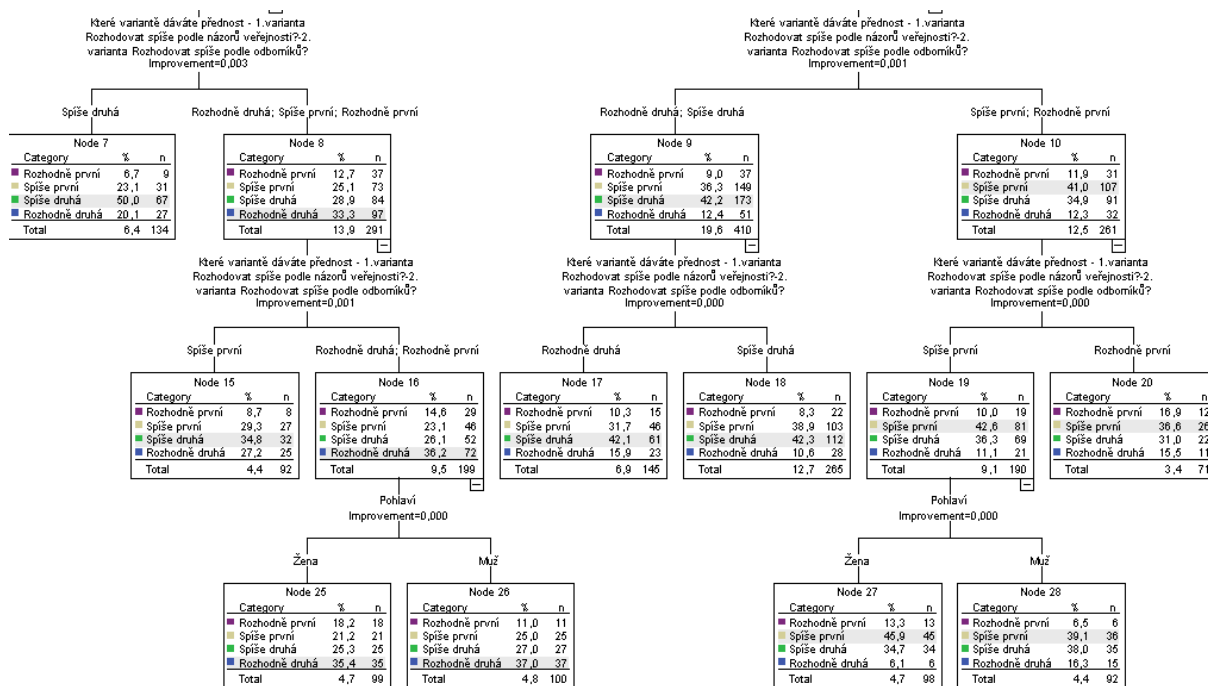


Zdroj: Vlastní práce, ČSDA

Levá větev se poté dále dělí podle další proměnné, tentokrát podle rozhodování dle názorů veřejnosti či odborných poradců. Až v poslední 5. úrovni, tak přichází „na řadu“ proměnná pohlaví. Osoby, které daly v prvních dvou úrovních přednost modernizaci, se pak ve větší části rozhodují stejně nehledě na další otázku. Zlom nastává v dalších úrovních, kde většina respondentů, kteří zvolil v první vysvětlující otázce možnost „rozhodně druhá“, pak na svém výběru vysvětlované proměnné trvají stále, zatímco u osob, volící střední možnost „spíše druhá“ je už ve čtvrté úrovni vidět zlom a jejich otočení

na životní prostředí. V páté úrovni je vidět už tolikrát řečená věta, že není možné podporovatele modernizace klasifikovat podle pohlaví.

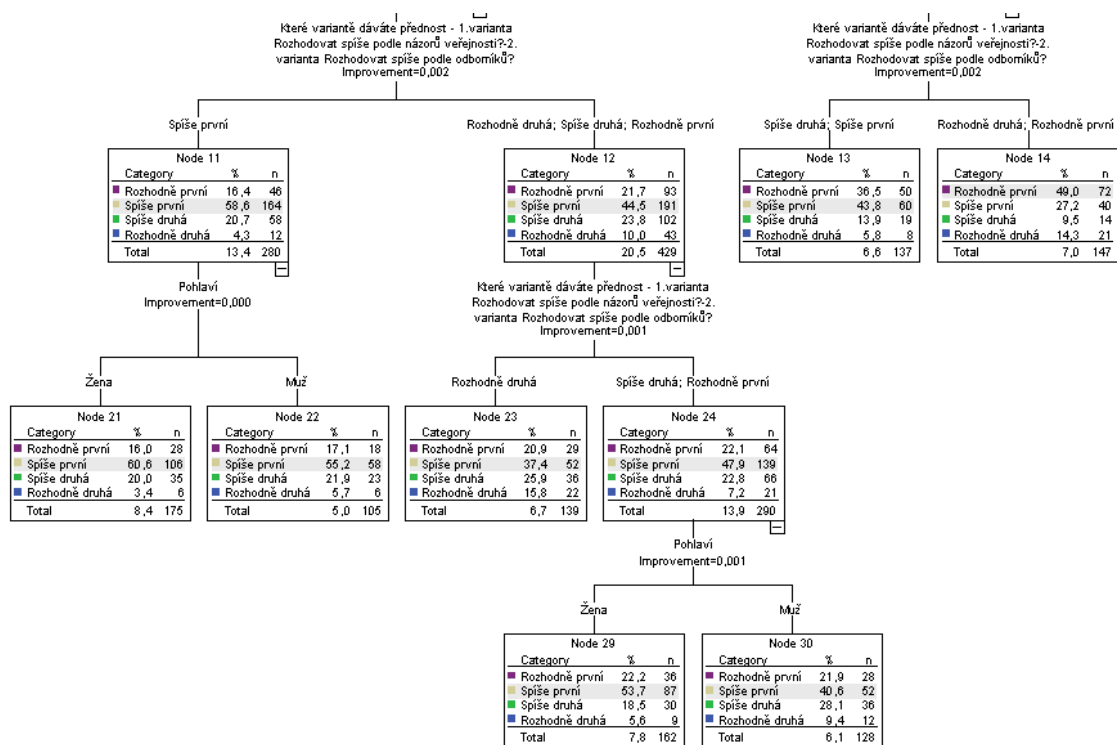
Obrázek 12: Výstup z programu IBM SPSS – Levá větev metody CART – analýza č.4



Zdroj: Vlastní práce, ČSDA

Na pravé větvi stromu, kde jsou osoby volící u první proměnné dodržování tradic je vidět ustálenější názor, kdy v podstatě až do páté úrovně trvají na péči o životní prostředí, nehledě na další vysvětlující proměnné.

Obrázek 13: Výstup z programu IBM SPSS – Pravá větev metody CART – analýza č.4



Zdroj: Vlastní práce, ČSDA

Klasifikace na základě výsledků metody dopadla o 12 % lépe než v předchozím případě pomocí diskriminační analýzy. Celková neúspěšnost odhadu je 57,3 % a naopak úspěšnost dosáhla hodnoty 42,7 %. Tento výsledek je přijatelný a tato analýza nám především ukázala stálost názoru lidí zastávajících se tradic, kteří neustále trvali na dodržování životního prostředí. Osoby, které však spíše zastávají názor modernizace už pak tuto názorová stálost nemají.

Tabulka 29: Výstup z programu IBM SPSS – Tabulka Risk – analýza č.4

Estimate	Std. Error
,573	,011

Growing Method:
CRT
Dependent Variable:
Jakými zásadami by se měla řídit společnost, v níž byste chtěl/a žít-1.
varianta Starat se hlavně o dobré životní prostředí?- 2.varianta Starat se hlavně o ekonomický růst?

Zdroj: Vlastní práce, ČSDA

Tabulka 30: Výstup z programu IBM SPSS – Tabulka Risk – analýza č.4

Observed	Predicted				Percent Correct
	Rozhodně první	Spíše první	Spíše druhá	Rozhodně druhá	
Rozhodně první	122	139	114	0	32,5%
Spíše první	100	355	360	0	43,6%
Spíše druhá	33	160	415	0	68,3%
Rozhodně druhá	29	55	207	0	0,0%
Overall Percentage	13,6%	33,9%	52,5%	0,0%	42,7%

Zdroj: Vlastní práce, ČSDA

4.10 Porovnání výsledků obou metod

Při porovnání výsledků obou metod je nejprve nutné si uvědomit, že každá z těchto dvou metody zastává naprosto rozdílný přístup ke klasifikaci.

Zatímco diskriminační analýza nejprve zjišťuje závislost mezi proměnnými, tak jak je popsáno v teoretické části práce, tak metoda CART pracuje na bázi Giniho indexu, pomocí jehož výsledků následně proměnné klasifikuje.

Metody je však možné porovnat v rámci toho, jak kvalitně dokázaly klasifikovat proměnné do správných skupin na základě výsledků jejich analýz.

V první analýze, týkající se přeměny společnosti spojenou s modernizací se diskriminační analýze podařilo správně zařadit 28,5 % proměnných. Z toho se nejlépe podařilo v krajních odpovědích na vysvětlovanou otázku. Analýza si na základě zjištěné závislosti nedokázala správně zařadit „prostřední“ proměnné. Metoda CART měla pozitivnější výsledek následné klasifikace a to sice 45,3 %. Zajímavé je vidět rozdíl mezi správným zařazením diskriminační analýzy a metoda CART. Metodě CART se totiž následná klasifikace podařila lépe ve výše zmíněných prostředních hodnotách, zatímco u krajních hodnot nebyla úspěšnost příliš vysoká.

V případě druhé analýzy, zaměřené na globalizaci v rámci modernizace, vyšel diskriminační analýze nejlepší výsledek ze všech čtyř provedených analýz. Klasifikovat proměnné se povedlo v 42,5 % případů. Nejlépe se to opět povedlo v krajních možnostech, kde se například možnost „rozhodně ano“ podařilo klasifikovat v 70,3 % případů. Výsledné procent u klasifikace však srážely střední možnosti, kde klasifikace nikdy nepřekročila 25 %. Metoda CART dokázala na základě Giniho indexu správně klasifikovat proměnné v 54,5 % a stejně jako v případě diskriminační analýzy se jednalo o nejlepší

výsledek ze všech čtyř analýz. Nejvyšší procentuální úspěšnost zaznamenal u možnosti „rozhodně ne“, kde byla úspěšnost vyšší než 80 %. Co však této klasifikaci uškodilo, byla naprosto 0 % úspěšnost klasifikace u jedné ze středních proměnných.

Třetí analýza zaměřující se na zjištění, zdali je možné, podle určitých typologických znaků, říci, které osoby jsou pro modernizaci a které nikoli, dopadla pro obě dvě metody nejhůře ze všech čtyř metod. Metoda CART měla v následné klasifikaci úspěch pouze 37,5 % a diskriminační analýza dokonce pouze v 26,6 % případů. Metodě CART se zde dařilo zařazovat lépe prostřední možnosti, ale opět ji srazila 0 % úspěšnost v jedné z krajních variant. Diskriminační analýza měla naopak úspěch v krajních možnostech u vysvětlované proměnné, ale velmi nízkou úspěšnost v možnostech středních.

Poslední analýza týkající se ekonomického růstu a péče o životní prostředí dopadla o něco lépe než předchozí analýza. Diskriminační analýza správně klasifikovala 30,5 % proměnných a metoda CART 42,7 %. Zatímco u diskriminační analýzy se podařilo „překonat“ výsledek z první analýzy, metoda CART zaznamenala výsledek o necelé 3 % horší.

V celkovém klasifikačním srovnání tedy metoda CART zvládla klasifikovat proměnné na základě předchozích výsledků v průměru o 13 % lépe než diskriminační analýza. Lepší klasifikaci zvládla ve všech čtyřech analýzách. Diskriminační analýza naopak nabídla i pohled na závislosti mezi jednotlivým proměnnými.

5 ZÁVĚR

V praktické části této práce byly provedeny 4 analýzy, nejprve pro diskriminační analýzu a poté pro metodu klasifikačních a regresních stromů.

V analýze přeměny společnosti na tempu modernizace se podařilo pomocí diskriminační analýzy nalézt slabou závislost, která však nemohla být náhodná, což potvrdil test Wilksova lambda rozdělení, a s její pomocí se následně podařilo správně klasifikovat 28,5 % proměnných z celkového souboru 2090 respondentů. Ve stejné analýze za pomoci metody CART se podařilo správně klasifikovat 45,3 % respondentů a metoda CART tedy zaznamenala o 16,8 % lepší výsledek než diskriminační analýza.

V druhé analýze zkoumající globalizaci v rámci modernizace se pomocí diskriminační analýzy podařilo najít silnější závislost než v předchozí analýze s tím, že největší závislost na vysvětlované proměnné měla proměnná týkající se odsunu určitých skupin na okraj společnosti. Na základě těchto výsledků je tedy možné říci, že osoby, které nestojí o modernizaci, kdyby vedla k přistěhovalecké vlně z východu, by zároveň nechtěly, aby se v rámci modernizace odsunula určitá skupina obyvatel na okraj společnosti. Celková klasifikace byla provedena s 42,5 % úspěšností. Metoda CART v tomto případě zaznamenala opět vyšší úspěšnost klasifikace než diskriminační analýza. Její úspěšnost dosáhla 54,5 %.

Analýza postojů k modernizaci, která proběhla jako třetí, nabídla pro obě metody nejslabší výsledky. U diskriminační analýzy se nepodařilo zjistit prakticky žádnou závislost a výsledek klasifikace byl 26,6 %. U metody CART byla výsledkem nejslabší klasifikace proměnných ze všech čtyř analýz. Klasifikaci se podařilo správně provést pouze u 37,5 % respondentů. Z této analýzy vyplynulo, že na základě základních znaků jako jsou pohlaví, rodinný stav a vzdělání, není možné určit, zda je osoba pro či proti modernizaci.

Poslední provedenou analýzou byla analýza postojů k ekonomickému růstu. V případě diskriminační analýzy vyšla závislost, která říká, že osoby, jež se vyslovily spíše pro ekonomický růst než pro péči o životní prostředí, jsou zároveň podporovatelé modernizace než zastánci dodržování tradic. Celková klasifikace, provedena na základě výsledků diskriminační analýzy, byla úspěšná z 30,5 % a úspěšnost tedy byla vyšší než v případě první analýzy. U metody CART byla tato klasifikace úspěšná z 42,7 %, což je naopak o necelé 3 % nižší než v případě první analýzy.

V celkovém součtu všech analýz a vyhodnocení následných klasifikací dopadla lépe metoda klasifikačních a regresních stromů, tedy metoda CART. Metoda CART zaznamenala v průměru o 13 % lepší výsledek klasifikace než diskriminační analýza. Pro tento typ analýzy je tedy možné označit jako vhodnější metodu CART. Metoda CART má také graficky přehlednější výstupy, jež by nemusely být pro amatéra takovou překážkou, jako například výstupy z diskriminační analýzy, jež vyžadují nadprůměrnou znalost statistiky. Důvodem horších výsledků diskriminační analýzy může být například velká složitost tématu modernizace, kdy například podle základních znaků není možné určit osoby vyznačující se souhlasem či odpor pro modernizaci.

Hlavním přínosem této práce je popsání a porovnání výsledků obou metod na stejném datovém souboru, které může sloužit jako podklad pro další studenty, kteří by se chtěli tématu klasifikačních a regresních stromů věnovat. Dalším přínosem je analýza modernizace společnosti, které může být, v dnešní, čím rychle se měnící společnosti, velice přínosné.

6 SEZNAM ZDROJŮ

KNIŽNÍ ZDROJE

1. BECK, U. *Riziková společnost: na cestě k jiné moderně*. Vyd. 1. Přeložil Otakar VOCHOČ. Praha: Sociologické nakladatelství, 2004. Post. ISBN 80-86429-32-6.
2. BERKA, P. *Dobývání znalostí z databází*. Praha: Academia, 2003. ISBN 80-200-1062-9.
3. BERRY, M. J. A. - Linoff, G. *Data mining techniques for marketing, sales, and customer relationship management*. Indianapolis: Wiley, 2004. ISBN 0-471-47064-3.
4. BRYCHTOVÁ, Š. *Globalizace a krize v souvislostech: Globalization and crisis in connection*. Scientific Papers of the University of Pardubice. Series D, Faculty of Economics & Administration. 20, 29, 5-13, Nov. 2013. ISSN: 1211555X
5. GIDDENS, A. a MÜLLER A.. *Důsledky modernity / Anthony Giddens ; [z anglického originálu .. přeložil Karel Müller]*. 2010. ISBN 9788074190353
6. HEBÁK, P. -- JAROŠOVÁ, E. -- PECÁKOVÁ, I. -- PLAŠIL, M. -- ŘEZANKOVÁ, H. -- VILIKUS, O. -- VLACH, P. *Statistické myšlení a nástroje analýzy dat*. 1. vyd. Praha: Informatorium, 2013. 877 s. ISBN 978-80-7333-105-4.
7. KELLER, J. *Teorie modernizace*. Praha: Slon, 2007. ISBN: 978-80-86429-66-3.
8. PECÁKOVÁ, I. *Statistika v terénních průzkumech*. 2. vyd. Praha: PROFESSIONAL PUBLISHING, 2011. 236 s. ISBN 978-80-7431-039-3.
9. SPSS INC. *SPSS statistical algorithms*. Chicago: SPSS, 1991. ISBN 0-918469-89-9.

INTERNETOVÉ ZDROJE

10. JARKOVSKÝ, J. – LITNERROVÁ, S. *Vícerozměrné statistické metody*. Přednáška, Brno, Masarykova Univerzita.

URL:<http://www.iba.muni.cz/esf/res/file/bimat-prednasky/vicerozmerne-statisticke-metody/VSM-08.pdf>

11. KINGSFORD, C., & SALZBERG, S. L. (2008). *What are decision trees? Nature Biotechnology*, 26(9), 1011-3.

URL:<http://dx.doi.org/10.1038/nbt0908-1011>

12. KNOBL, W. "Theories That Won't Pass Away: The Never-ending Story." in Gerard Delanty, Engin F. Isin, eds. *Handbook of Historical Sociology* (2003): 96-107

URL:

[http://193.231.20.29/seminarii/locuire/Upload/syllabus09/C09b.%20Knobl%20\(2003\).pdf](http://193.231.20.29/seminarii/locuire/Upload/syllabus09/C09b.%20Knobl%20(2003).pdf)

13. MRÁZOVÁ, Iveta. *Dobývání znalostí*. Přednáška, Praha, UK-MFF.

URL:http://ksvi.mff.cuni.cz/~mraz/datamining/lecture/Dobyvani_Znalosti_Prednaska_Uvod.pdf

14. PLCHÚT, Martin. *Dobývání znalostí z databází: Úvod a oblasti aplikací*. In: euromise.vse.cz.

URL:<http://www.fit.vutbr.cz/study/courses/ZZD/public/seminar0304/Uvod.pdf>

15. SPEYBROECK, N. (2012). Classification and regression trees. *International Journal of Public Health*, 57(1), 243-246.

URL:<http://dx.doi.org/10.1007/s00038-011-0315-z>

16. KUMAR, K. *Modernization*. Encyclopedia Britannica. Britannica.com. Retrieved on 2013-08-17

URL: <https://www.britannica.com/topic/modernization>

7 Přílohy

Tabulka 31: Výstup z programu IBM SPSS – Rozložení respondentů v analýze č.1

Do jaké míry souhlasíte s následujícími výroky - Tempo modernizace západních společností je nesmyslně vysoké?		Valid N (listwise)	
		Unweighted	Weighted
Rozhodně souhlasím	Do jaké míry souhlasíte s výrokem - Naše společnost se mění tak rychle, že to člověk ani nestačí sledovat?	325	325,000
	Ohlédnete-li se zpět, myslíte si, že změna společenského systému u nás, k níž došlo v roce 1989, stála za to?	325	325,000
	Vy sám/sama jste ochoten/na v případě potřeby zvýšit své pracovní tempo?	325	325,000
Spíše souhlasím	Do jaké míry souhlasíte s výrokem - Naše společnost se mění tak rychle, že to člověk ani nestačí sledovat?	722	722,000
	Ohlédnete-li se zpět, myslíte si, že změna společenského systému u nás, k níž došlo v roce 1989, stála za to?	722	722,000
	Vy sám/sama jste ochoten/na v případě potřeby zvýšit své pracovní tempo?	722	722,000
Spíše nesouhlasím	Do jaké míry souhlasíte s výrokem - Naše společnost se mění tak rychle, že to člověk ani nestačí sledovat?	713	713,000
	Ohlédnete-li se zpět, myslíte si, že změna společenského systému u nás, k níž došlo v roce 1989, stála za to?	713	713,000
	Vy sám/sama jste ochoten/na v případě potřeby zvýšit své pracovní tempo?	713	713,000
Rozhodně nesouhlasím	Do jaké míry souhlasíte s výrokem - Naše společnost se mění tak rychle, že to člověk ani nestačí sledovat?	330	330,000
	Ohlédnete-li se zpět, myslíte si, že změna společenského systému u nás, k níž došlo v roce 1989, stála za to?	330	330,000
	Vy sám/sama jste ochoten/na v případě potřeby zvýšit své pracovní tempo?	330	330,000
Total	Do jaké míry souhlasíte s výrokem - Naše společnost se mění tak rychle, že to člověk ani nestačí sledovat?	2090	2090,000
	Ohlédnete-li se zpět, myslíte si, že změna společenského systému u nás, k níž došlo v roce 1989, stála za to?	2090	2090,000
	Vy sám/sama jste ochoten/na v případě potřeby zvýšit své pracovní tempo?	2090	2090,000

Zdroj: Vlastní práce, ČSDA

Tabulka 32: Výstup z programu IBM SPSS – Rozložení respondentů v analýze č.2

Group Statistics		Valid N (listwise)	
		Unweighted	Weighted
rozhodně ano	Prosím, řekněte mi, zda jste pro modernizaci, i kdyby vedla k silně vlně přistěhovalectví z východu?	118	118,000
	Prosím, řekněte mi, zda jste pro modernizaci, i kdyby vedla ke stěhování lidí za prací?	118	118,000
	Prosím, řekněte mi, zda jste pro modernizaci, i kdyby vedla k vytlačování některých skupin lidí na okraj společnosti? Vy sám/sama jste ochoten/na v případě potřeby odstěhovat se za prací do jiného města?	118	118,000
spíše ano	Prosím, řekněte mi, zda jste pro modernizaci, i kdyby vedla ke stěhování lidí za prací?	289	289,000
	Prosím, řekněte mi, zda jste pro modernizaci, i kdyby vedla k vytlačování některých skupin lidí na okraj společnosti?	289	289,000
	Prosím, řekněte mi, zda jste pro modernizaci, i kdyby vedla k vytlačování některých skupin lidí na okraj společnosti? Vy sám/sama jste ochoten/na v případě potřeby odstěhovat se za prací do jiného města?	289	289,000
spíše ne	Prosím, řekněte mi, zda jste pro modernizaci, i kdyby vedla ke stěhování lidí za prací?	697	697,000
	Prosím, řekněte mi, zda jste pro modernizaci, i kdyby vedla k vytlačování některých skupin lidí na okraj společnosti?	697	697,000
	Prosím, řekněte mi, zda jste pro modernizaci, i kdyby vedla k vytlačování některých skupin lidí na okraj společnosti? Vy sám/sama jste ochoten/na v případě potřeby odstěhovat se za prací do jiného města?	697	697,000
rozhodně ne	Prosím, řekněte mi, zda jste pro modernizaci, i kdyby vedla ke stěhování lidí za prací?	1122	1122,000
	Prosím, řekněte mi, zda jste pro modernizaci, i kdyby vedla k vytlačování některých skupin lidí na okraj společnosti?	1122	1122,000
	Prosím, řekněte mi, zda jste pro modernizaci, i kdyby vedla k vytlačování některých skupin lidí na okraj společnosti? Vy sám/sama jste ochoten/na v případě potřeby odstěhovat se za prací do jiného města?	1122	1122,000
Total	Prosím, řekněte mi, zda jste pro modernizaci, i kdyby vedla ke stěhování lidí za prací?	2226	2226,000
	Prosím, řekněte mi, zda jste pro modernizaci, i kdyby vedla k vytlačování některých skupin lidí na okraj společnosti?	2226	2226,000
	Prosím, řekněte mi, zda jste pro modernizaci, i kdyby vedla k vytlačování některých skupin lidí na okraj společnosti? Vy sám/sama jste ochoten/na v případě potřeby odstěhovat se za prací do jiného města?	2226	2226,000

Zdroj: Vlastní práce, ČSDA

Tabulka 33: Výstup z programu IBM SPSS – Rozložení respondentů v analýze č.3

Group Statistics

Jakými zásadami by se měla řídit společnost, v níž byste chtěl/a žít-1.varianta Upřednostňovat dodržování tradic?-2.varianta Upřednostňovat modernizační změny?		Valid N (listwise)	
		Unweighted	Weighted
Rozhodně první	Jaké je vaše nejvyšší ukončené školní vzdělání?	304	304,000
	Pohlaví	304	304,000
	Rodinný stav	304	304,000
Spíše první	Jaké je vaše nejvyšší ukončené školní vzdělání?	766	766,000
	Pohlaví	766	766,000
	Rodinný stav	766	766,000
Spíše druhá	Jaké je vaše nejvyšší ukončené školní vzdělání?	728	728,000
	Pohlaví	728	728,000
	Rodinný stav	728	728,000
Rozhodně druhá	Jaké je vaše nejvyšší ukončené školní vzdělání?	462	462,000
	Pohlaví	462	462,000
	Rodinný stav	462	462,000
Total	Jaké je vaše nejvyšší ukončené školní vzdělání?	2260	2260,000
	Pohlaví	2260	2260,000
	Rodinný stav	2260	2260,000

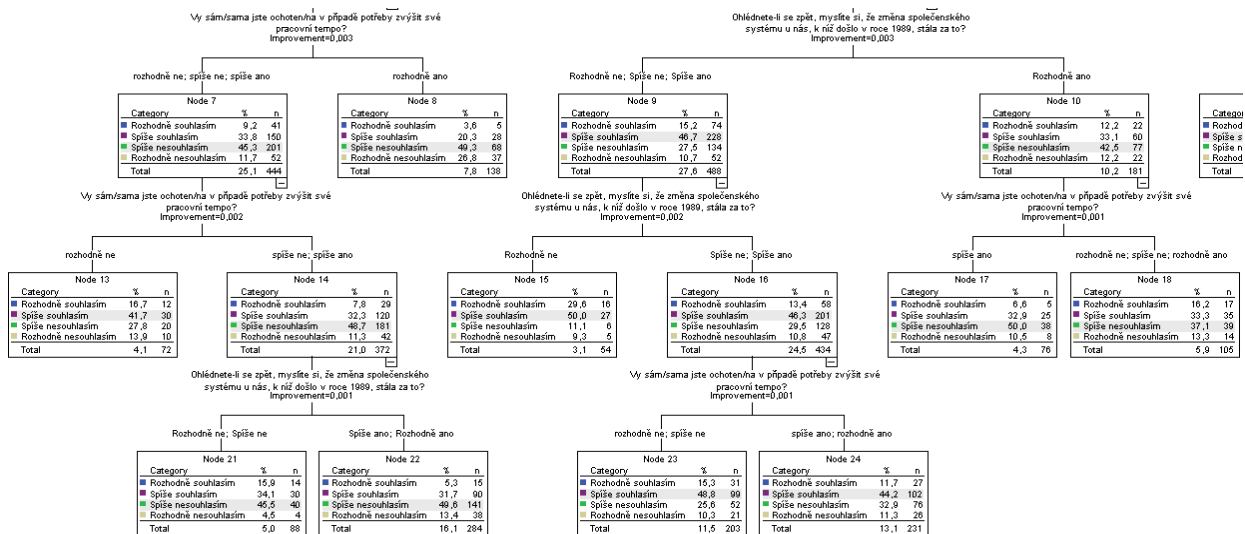
Zdroj: Vlastní práce, ČSDA

Tabulka 34: Výstup z programu IBM SPSS Rozložení respondentů v analýze č.4

Jakými zásadami by se měla řídit společnost, v níž byste chtěl/a žít-1.varianta Starat se hlavně o dobré životní prostředí?- 2.varianta Starat se hlavně o ekonomický růst?		Valid N (listwise)	
		Unweighted	Weighted
Rozhodně první	Pohlaví Jakými zásadami by se měla řídit společnost, v níž byste chtěl/a žít-1. varianta Upřednostňovat dodržování tradic?-2. varianta Upřednostňovat modernizační změny? Které variantě dáváte přednost - 1.varianta Rozhodovat spíše podle názorů veřejnosti?-2. varianta Rozhodovat spíše podle odborníků?	409 409 409	409,000 409,000 409,000
Spíše první	Pohlaví Jakými zásadami by se měla řídit společnost, v níž byste chtěl/a žít-1. varianta Upřednostňovat dodržování tradic?-2. varianta Upřednostňovat modernizační změny? Které variantě dáváte přednost - 1.varianta Rozhodovat spíše podle názorů veřejnosti?-2. varianta Rozhodovat spíše podle odborníků?	867 867 867	867,000 867,000 867,000
Spíše druhá	Pohlaví Jakými zásadami by se měla řídit společnost, v níž byste chtěl/a žít-1. varianta Upřednostňovat dodržování tradic?-2. varianta Upřednostňovat modernizační změny? Které variantě dáváte přednost - 1.varianta Rozhodovat spíše podle názorů veřejnosti?-2. varianta Rozhodovat spíše podle odborníků?	669 669 669	669,000 669,000 669,000
Rozhodně druhá	Pohlaví Jakými zásadami by se měla řídit společnost, v níž byste chtěl/a žít-1. varianta Upřednostňovat dodržování tradic?-2. varianta Upřednostňovat modernizační změny? Které variantě dáváte přednost - 1.varianta Rozhodovat spíše podle názorů veřejnosti?-2. varianta Rozhodovat spíše podle odborníků?	317 317 317	317,000 317,000 317,000
Total	Pohlaví Jakými zásadami by se měla řídit společnost, v níž byste chtěl/a žít-1. varianta Upřednostňovat dodržování tradic?-2. varianta Upřednostňovat modernizační změny? Které variantě dáváte přednost - 1.varianta Rozhodovat spíše podle názorů veřejnosti?-2. varianta Rozhodovat spíše podle odborníků?	2262 2262 2262	2262,000 2262,000 2262,000

Zdroj: Vlastní práce, ČSDA

Obrázek 14: Výstup z programu IBM SPSS Lev a prostřední větve – analýza č. 1



Zdroj: Vlastní práce, ČSDA