

**Univerzita Pardubice**  
**Fakulta ekonomicko-správní**

Návrh klasifikátora parametrů udržitelného rozvoje pomocí  
RBF neuronových sítí  
Jitka Skřivánková

Diplomová práce  
2010

Univerzita Pardubice  
Fakulta ekonomicko-správní  
Ústav systémového inženýrství a informatiky  
Akademický rok: 2009/2010

## ZADÁNÍ DIPLOMOVÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Jitka SKŘIVÁNKOVÁ**  
Studijní program: **N6209 Systémové inženýrství a informatika**  
Studijní obor: **Informatika ve veřejné správě**

Název tématu: **Návrh klasifikátora parametrů udržitelného rozvoje pomocí RBF neuronových sítí**

### Z á s a d y p r o v y p r a c o v á n í :

Charakterizujte udržitelný rozvoj.  
Analyzujte vstupní data (parametry) pro následující klasifikaci.  
Charakterizujte RBF neuronové sítě z hlediska klasifikace a aproximace.  
Navrhněte model na modelování udržitelného rozvoje.  
Verifikujte navržený model.  
Uskutečňte analýzu výsledků.

Rozsah grafických prací:

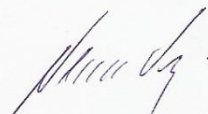
Rozsah pracovní zprávy:

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

Seznam odborné literatury:

1. OLEJ, V.: Modelovanie ekonomických procesov na báze výpočtovej inteligencie. [Vedecká monografia], Miloš Vognar - M&V, ISBN 80-903024-9-1, Hradec Králové, Česká republika, 2003, 160s.
2. KVASNIČKA, V. a kol.: Úvod do teórie neurónových sietí. Iris, Bratislava, 1997.
3. HAYKIN, S.: Neural Networks: A Comprehensive Foundation. 2nd edition, New Jersey, Prentice-Hall, Inc., 1999, 842s.
4. FIALOVÁ, H. a kol.: Vybrané oblasti udržitelného rozvoje v Pardubickém kraji. Český statistický úřad, Oddelení regionálních analýz a informačních služeb Pardubice, Pardubice, 2007.

Vedoucí diplomové práce:


  
**prof. Ing. Vladimír Olej, CSc.**  
Ústav systémového inženýrství a informatiky

Datum zadání diplomové práce:

**5. října 2009**

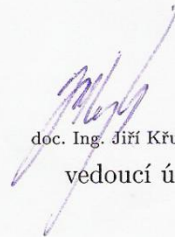
Termín odevzdání diplomové práce:

**30. dubna 2010**

  
doc. Ing. Renáta Myšková, Ph.D.

děkanka

L.S.

  
doc. Ing. Jiří Křupka, Ph.D.

vedoucí ústavu

V Pardubicích dne 5. října 2009

# Prohlášení

Tuto práci jsem vypracovala samostatně. Veškeré literární prameny a informace, které jsem v práci využila, jsou uvedeny v seznamu použité literatury.

Byla jsem seznámena s tím, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorský zákon, zejména se skutečností, že Univerzita Pardubice má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60 odst. 1 autorského zákona a s tím, že pokud dojde k užití této práce mnou nebo bude poskytnuta licence o užití jinému subjektu, je Univerzita Pardubice oprávněna ode mne požadovat přiměřený příspěvek na úhradu nákladů, které na vytvoření díla vynaložila, a to podle okolností až do jejich skutečné výše.

Souhlasím s prezenčním zpřístupněním své práce v Univerzitní knihovně Univerzity Pardubice.

V Pardubicích dne 5.4.2010

Jitka Skřivánková

## **Poděkování**

Ráda bych touto cestou poděkovala vedoucímu diplomové práce panu prof. Ing. Vladimíru Olejovi, CSc. za odbornou pomoc, připomínky a jeho podporu při vypracovávání této diplomové práce.

Dále děkuji svým rodičům za jejich pochopení, trpělivost a také za to, že mi umožnili studovat.

## **Souhrn**

Diplomová práce se zabývá návrhem parametrů udržitelného rozvoje pro klasifikaci. Pro klasifikaci byla vybraná metoda neuronové sítě typu radiálně bazických funkcí. Obce Pardubického kraje byly klasifikovány do tříd podle definovaných kritérií. Cílem práce jsou návrhy modelů udržitelného rozvoje na základě tří pilířů, analýza modelů a jejich verifikace.

## **Klíčová slova**

Udržitelný rozvoj, indikátory, RBF neuronová síť, klasifikace.

## **Tittle**

Classifier of parameters sustainable development design by RBF neural network.

## **Abstract**

This thesis deals with design parameters for the classification of sustainable development. The method of neural network-type radial basis functions was selected for this classification. Municipalities of Pardubice Region were classified into classes according to defined criteria. The aim of the proposed model work for sustainable development based on three basic pillars, analyze models and their verification.

## **Keywords**

Sustainable development, parameters, RBF neural network, classification.

# Obsah

Úvod.....	9
1 Udržitelný rozvoj .....	10
1.1 Strategie udržitelného rozvoje .....	12
1.1.1 Strategie udržitelného rozvoje EU .....	12
1.1.2 Strategie v České republice.....	13
1.2 Indikátory udržitelného rozvoje.....	13
1.3 Návrh indikátorů .....	14
2 RBF neuronová síť .....	16
2.1 Neuronové sítě .....	16
2.2 Topologie RBF neuronové sítě .....	19
2.3 Aktivační funkce RBF neuronové sítě.....	20
2.4 Učení RBF neuronové sítě .....	21
2.4.1 Učení s učitelem.....	21
2.4.2 Počet neuronů ve skryté vrstvě .....	22
2.4.3 Nalezení center RBF .....	24
2.5 Všeobecný klasifikační problém.....	25
2.6 Klasifikace .....	27
3 Návrh modelu udržitelného rozvoje.....	28
3.1 Analýza parametrů .....	29
3.1.1 Ekonomický pilíř.....	29
3.1.2 Environmentální pilíř .....	31
3.1.3 Sociální pilíř .....	32
3.2 Předzpracování dat.....	35
3.3 Návrh struktury RBF neuronové sítě .....	37

4	Analýza výsledků navržené struktury RBF.....	39
4.1	Ekonomický model .....	39
4.2	Environmentální model.....	44
4.3	Sociální model .....	48
4.4	Výsledky klasifikace navržených modelů .....	53
5	Aproximace .....	54
5.1	Ekonomický model .....	55
5.2	Environmentální model.....	56
5.3	Sociální model .....	62
5.4	Výsledky aproximace navržených modelů .....	67
	Závěr .....	69
	Seznam obrázků.....	70
	Seznam tabulek .....	71
	Seznam grafů .....	72
	Seznam příloh .....	73
	Použitá literatura .....	74



# Úvod

Udržitelný rozvoj se stal pojmem, který se v poslední době dostává do popředí. Co si vlastně představit pod pojmem udržitelný rozvoj, vypovídá první část diplomové práce. V této části je uvedena se strategií udržitelného rozvoje Evropské unie (EU) a České republiky (ČR). Tato kapitola je věnována indikátorům udržitelného rozvoje.

Druhá část seznamuje s Radiál Basic Function (RBF) neuronovou sítí, která nám umožní klasifikovat obce podle parametrů udržitelného rozvoje.

O jednotlivých parametrech, které byly navrhnuty expertem je pojednáno ve třetí kapitole. Zde je také navržen model pro návrh klasifikátoru parametru udržitelného rozvoje pomocí RBF.

Ve čtvrté části je podrobná analýza modelů udržitelného rozvoje. V poslední části je provedena aproximace, na jejímž základě je pak návrh na vyřazení některých parametrů udržitelného rozvoje.

Cílem diplomové práce je nalezení vhodných parametrů pro klasifikaci udržitelného rozvoje a návrh modelu udržitelného rozvoje pro klasifikaci obcí. Verifikace modelu je uskutečněna v prostředí Clementine.

# 1 Udržitelný rozvoj

Všechny organismy na planetě ovlivňují své okolí, také člověk je jedním z těchto narušitelů. Na počátku narušoval člověk okolní prostředí jen minimálně, žil jako součást přirozeného ekosystému a jeho vlivy byly značně omezené. V té době byly ohroženy některé druhy velkých savců, jako např. mamuti. Velký zlom pro člověka nastal koncem doby ledové, člověk se díky příznivým klimatickým podmínkám začal stěhovat do dosud neobydlených severních oblastí. Lidská společnost se tak začala věnovat zemědělství a pastevectví a přeměňovala okolní ráz krajiny na pastviny a pole. Další významnou změnou byla průmyslová revoluce, která znamenala nástup složitých strojů a počátek využívání fosilních paliv. To umožnilo člověku s mnohem větší intenzitou ovlivňovat okolní prostředí. Poslední období vývoje lidské společnosti je industriální, které umožňuje člověku působit trvalé či dlouhodobé změny v globálním rozsahu. S růstem schopností člověka využívat energii a přírodní zdroje pro uspokojení svých potřeb, požadavků i tužeb a s rostoucím počtem jedinců, kteří chtějí tyto potřeby, požadavky a tužby realizovat, roste tlak na prostředí. Za hlavní příčiny ovlivňující životní prostředí na Zemi lze označit [17]:

- Stále se zvyšující nároky na uspokojování lidských potřeb.
- Růst lidské populace.

Položme si tedy otázku: „Mohou tyto změny ovlivnit nějak život následujících generací a ohrozit životní prostředí?“ Vždyť ne všechny zdroje jsou obnovitelné. Tento problém si začalo lidstvo uvědomovat v 70. letech minulého století. Mezi první knihy, které se začaly věnovat této tématice, patří **Meze růstu** od D. Meadows (The Limits to growth. Univerce Books, New York 1972). Tato kniha šokovala svět thezí, že nejspíše do roku 2100 planeta Země narazí na meze růstu. Pokud by se nezastavil neomezený průmyslový růst, dokonce je překročí [17].

V témže roce, kdy byla vydána kniha Meze růstu, se konala první konference ve Stockholmu, kde se došlo k závěru, že světu hrozí globální ekologická krize. Světová společnost byla vyzvána k akcím na ochranu životních prostředí: vzduch, voda, půda a živé přírody. Jako hlavní viník byl označen hospodářský růst, který je spojen stále s rostoucí průmyslovou výrobou, s vyšší spotřebou energie a všech materiálů, s rozšiřováním dopravy a s růstem měst. To sebou nese rychlé vyčerpání přírodních

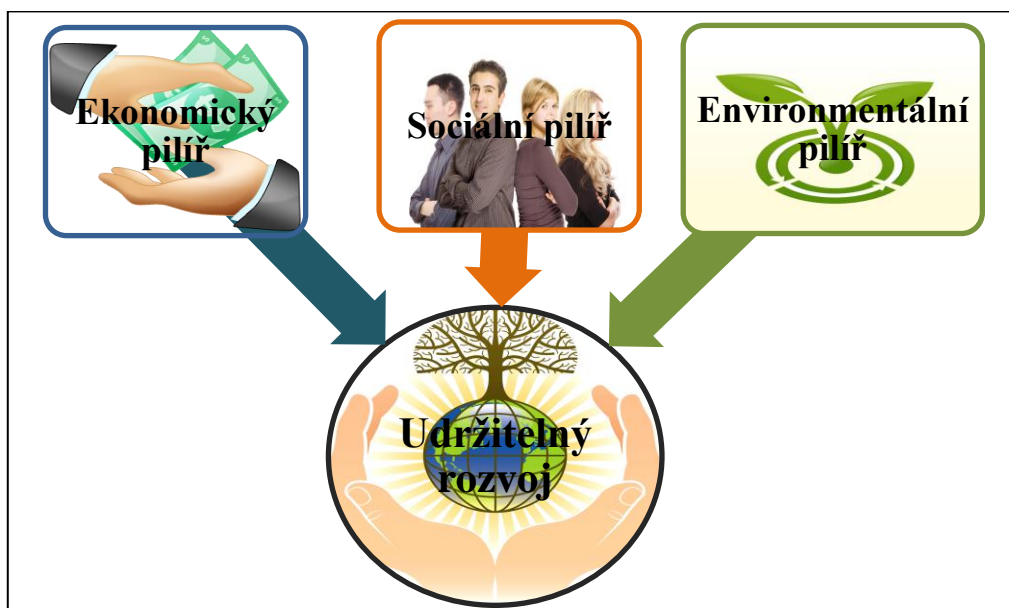
zdrojů, které se postupně mění na odpad. Někteří proto v 70. letech navrhovali, aby byl zastaven hospodářský růst [3], [19].

Na tuto tematiku navázala v roce 1987 zpráva „**Naše společná budoucnost**“, která byla vypracována pod vedením norské předsedkyně vlády Gro Harlem Bruntland. Zpráva definuje pojem udržitelný rozvoj. Podle zprávy je „Trvale udržitelný rozvoj takový způsob, který uspokojuje potřeby přítomnosti, aniž by oslaboval možnosti budoucích generací naplňovat jejich vlastní potřeby“. Přijetí myšlenky trvale udržitelného rozvoje bylo na konferenci v Riu r. 1992 [17], [18], [19].

Další definice udržitelného rozvoje je podle zákona České republiky č. 17/Sb. z roku 1992: „Trvale udržitelný rozvoj společnosti je takový rozvoj, který současné i budoucím generacím zachovává možnosti uspokojovat jejich základní životní potřeby a přitom nesnižovat rozmanitost přírody a zachovává přirozené funkce ekosystémů“. Od pojmu „trvale udržitelný rozvoj“ se v současné době opouští a používá se jen „udržitelný rozvoj“, význam je ale pořád stejný [20].

Udržitelný rozvoj se stal novým rámcem strategie civilizačního rozvoje. Cílem tohoto rozvoje je zajistit rovnováhu mezi třemi základními pilíři, to bylo zdůrazněno na světovém summitu v Johannesburgu v roce 2002. Základními pilíři udržitelného rozvoje jsou (viz. obrázek 1) [6]:

- ekonomický pilíř;
- environmentální pilíř;
- sociální pilíř.



Obrázek 1: Základní pilíře udržitelného rozvoje [zdroj: vlastní]

Podstatou udržitelného rozvoje je naplnění tří základních cílů [6]:

- sociální rozvoj, který respektuje potřeby všech;
- účinná ochrana životního prostředí a šetrné využívání přírodních zdrojů;
- udržení vysoké a stabilní úrovně ekonomického růstu a zaměstnanosti.

## **1.1 Strategie udržitelného rozvoje**

Strategie udržitelného rozvoje je dokument, který vymezuje hlavní strategické cíle k dosažení udržitelného rozvoje i způsoby a cesty. Ty umožňují tyto cíle naplnit. Strategie se opírá o řadu principů, z nichž nejdůležitější je úcta k lidskému životu a k přírodě, k civilizačním a kulturním hodnotám. Dokument je rozdělen do 12 oblastí, které rozvíjejí původních šest oblastí. Jedná se o ekonomický pilíř, environmentální pilíř, sociální pilíř a dále o následující oblasti: výzkum, vývoj a vzdělávání, evropský a mezinárodní kontext a správa věcí veřejných. Národní strategie udržitelného rozvoje zavádí Agendu 21 na úroveň státu, přitom zohledňuje specifické podmínky dané země [3], [23].

### **1.1.1 Strategie udržitelného rozvoje EU**

První Strategie udržitelného rozvoje EU byla přijata v roce 2001, na ni navázala v roce 2006 nová Strategie EU. Strategie reaguje na nadále přetrvávající trendy, zejména týkající se změn klimatu a spotřeb. Bohužel se jí nepodařilo zvrátit ani negativní jevy, jako např. hrozby chudoby a sociálního vyloučení, stárnutí populace. Cíle první strategie, které platí nadále, jsou [3]:

- Minimalizování změn klimatu a nákladů, které vyvolávají negativní efekty na společnost a životní prostředí. Zaměřuje se na zvyšování podílu energie vyrobené z obnovitelných zdrojů a podílu zelených paliv.
- Druhý cíl strategie se zaměřuje na to, aby dopravní systémy vyhovovaly ekonomickým a sociálním potřebám společnosti a přitom minimálně zatěžovaly jak společnost a ekonomiku, tak i životní prostředí.
- Třetí oblast se zabývá udržitelnou spotřebou a výrobou. Jde především o podporu udržitelných spotřebních zvyklostí a výrobních metod.
- Obsahem čtvrté oblasti je šetrné nakládání s přírodními zdroji a ochrana před jejich nadměrným užíváním. To se týká především lovu ryb, vody, půdy, vzduchu a atmosféry.

- Veřejné zdraví je pátou oblastí, jeho smyslem je podporovat veřejné zdravotnictví založené na rovných podmínkách pro všechny a na lepší ochraně před možnými zdravotními ohroženími. S tím souvisí i přiměřená prevence zdraví, zlepšení legislativy.
- Předposlední oblast tvoří sociální začlenění, demografie a migrace. Základním nástrojem je podpora solidarity mezi generacemi a v rámci generací, úsilí o vyšší kvalitu života občanů, která je základní podmínkou pro blahobyt jednotlivce.
- Globální chudoba a výzvy udržitelného rozvoje je poslední oblastí. Cílem je aktivně podporovat udržitelný rozvoj v celosvětovém měřítku a zajistit, aby vnitřní a vnější politika EU byla v souladu s globálním udržitelným rozvojem a s jejími mezinárodními závazky.

### **1.1.2 Strategie v České republice**

Strategie udržitelného rozvoje ČR je obdobná jako Strategie EU. Vymezuje prioritní oblasti, hlavní cíle a nástroje na jejich dosažení. Základním cílem zůstává minimalizovat prostor pro vznik nerovnováhy mezi ekonomickým, sociálním a environmentálním pilířem, což se považuje za základ jejich optimálního vývoje. Dalším úkolem je upozorňovat na existující a potenciální překážky a navrhnout vhodná opatření k jejich zmírnění, jedná-li se o negativní dopady. Tento dokument by se měl stát východiskem pro strategická rozhodnutí vlády a ministerstev [3].

V roce 2006 byla ke strategii udržitelného rozvoje sepsána situační zpráva, jež má za cíl podle zvolených indikátorů kontrolovat vývoj v České republice s ohledem na vytýčené cíle. Na základě sady 34 indikátorů popisuje vývoj ve třech pilířích udržitelného rozvoje a ve třech dalších významných oblastech [3].

## **1.2 Indikátory udržitelného rozvoje**

K lepšímu pochopení udržitelného rozvoje se používají indikátory. První zmínky o indikátorech vztahujících se k udržitelnému rozvoji byly v roce 1992 na konferenci v Riu, kde byla přijata Agenda 21. V této Agendě některé kapitoly vyjadřují nutnost vytvoření indikátorů, jako adekvátního podkladu pro rozhodování. První návrhy indikátorů se začaly objevovat již před více než 40 lety v USA. Obtížnost vystihnout všechny aspekty udržitelného rozvoje v jeden indikátor vedlo ke vzniku celé řady různých indikátorů. Existuje tak škála indikátorů od dílčích, zaměřených na určitou složku ekonomickou, životní prostředí nebo složku sociální. Indikátory vypovídají

o míře udržitelnosti či neudržitelnosti rozvoje. Indikátory slouží nejen k hodnocení úspěšností strategie, ale i k definování obsahu pojmu udržitelný rozvoj. Umožňují měřit a vyhodnocovat jednotlivé základní složky udržitelnosti. Měření základních složek udržitelnosti je velmi obtížné. Mezi základní příčiny obtížnosti měření a priority udržitelnosti jsou [17], [18]:

- Nejasná vazba mezi environmentálním, ekonomickým a sociálním pilířem. V teoretické rovině platí, že tyto tři pilíře by měly být ve vzájemné rovnováze. V praxi se však dostávají do rozporu např. ochrana přírody versus výstavba dálnice.
- Nedostatek dat o některých základních parametrech udržitelnosti. Nedostatek základních dat může znemožňovat vytváření relevantních indikátorů. V některých případech je obtížné porovnávat data z různých míst, či časových období. Řada klíčových hodnot, které jsou spojována s udržitelným rozvojem nelze převádět do měřitelných údajů, např. kvalitu života.
- Přebytek dílčích, specializovaných dat a informací. Na jedné straně existuje ohromné a narůstající množství nejrůznějších expertiz, studií a dílčích dat, srozumitelných pouze úzkému okruhu specialistů a na straně druhé stojí požadavky na poučené a fundované plánování a rozhodování o udržitelném rozvoji. Úspěšná strategie udržitelného rozvoje by se také měla stát věcí každého občana.

### 1.3 Návrh indikátorů

Kvůli použitelnosti indikátorů by měly indikátory splňovat několik požadavků, z nichž ty následující lze považovat za nejdůležitější [18], [27]:

- Reprezentativnost: musí být zřejmé, jaký předmět nebo jev, daný indikátor reprezentuje. Kromě volby správného obsahu je nutné zvolení vhodného geografického měřítka, případně vhodné časové rozložení měření či odebrání vzorků, jejichž analýzy jsou podkladem pro indikátory. Jestliže je například předmětem zájmu kvalita vody v nějaké řece, musí být jasné, kde a kdy se vzorky mají odebrat. Stejně tak je důležité, která měření a které analýzy se mají provádět.

- Spolehlivost: data musí být prověřována co do své spolehlivosti, potvrzována několika nezávislými měřeními, případně výsledky získanými zásadně různými metodami. Kontrola a zajištění kvality dat je důležitým oborem.
- Pochopitelnost: veškerá data předpokládají vždy nějakého uživatele, nějakého zájemce. Předpokladem jakéhokoliv využití údajů je jejich jasná pochopitelnost, jednoznačnost a srozumitelná prezentace.
- Jedinečnost: získané údaje mají být jedinečné, nemají být redundantní, opakované, dublovat nějaké již existující informace. Každý indikátor má mít svou specifičnost a originalitu a nesmí opakovat to, co již je známo odjinud.
- Náklady a užitek: pořízení, zpracování a poskytování jakýchkoliv informací stojí vždycky nějaké náklady. Pořizování dat, provoz monitorovacích systémů a provoz informačních systémů je obvykle záležitostí velmi nákladnou. V mnoha případech jsou tyto náklady nesrovnatelné s užitky. Dalším požadavkem je, aby byly vynaložené náklady úměrné k jejich užitku.
- Reálná zjistitelnost: zjištění musí být metodicky a technicky možné a to buď vlastním šetřením, nebo převzetím z existujících statistických databází.

## 2 RBF neuronová síť

Tato kapitola se zaměřuje na Radial Basis Function (RBF) neuronové síť. RBF se poprvé objevila v 80. letech. Jedná se o dopřednou neuronovou síť, jejíž aktivační funkce využívá radiální bázovou funkci. Z aproximačního hlediska jsou takto vymezené neuronové síťe přirozené, a to z toho důvodu, že jsou aproximovány funkcemi, jež neovlivní výslednou funkci v celém rozsahu, ale jen okolí centra příslušného neuronu RBF[1], [5].

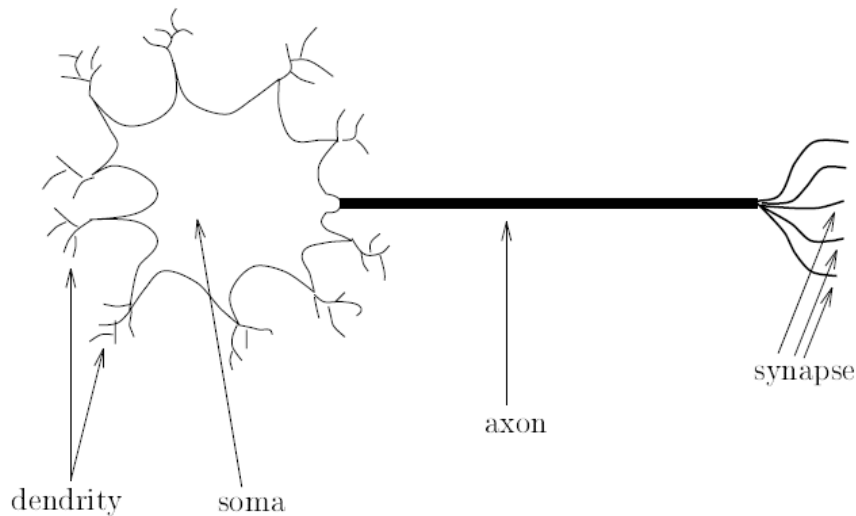
### 2.1 Neuronové síťe

Co to ale vlastně ta neuronová síť je? Umělé neuronové síťe (UNS) jsou jedním ze směrů výzkumu v oblasti umělé inteligence. Svoji existenci založila umělá inteligence na teorii, že lze napodobit inteligenci živých organismů, zejména lidského mozku. Díky činnosti nervové soustavy a mozku je člověk schopen vykonávat mnoho činností, které ovlivňují kvalitu jeho života. Nervový systém je nejsložitější síť v těle, má rozhodující význam pro rozvoj řeči, myšlení a paměti. Centrem nervového systému jsou mozek a mícha, které kontrolují všechny nervové tkáně v ostatních částech těla. Aktivními částmi nervového systému jsou miliony vzájemně propojených nervových buněk nazývaných neurony. Neurony mají podobnou funkci jako drátky ve složitém elektronickém přístroji. Přijímají vzruchy z jedné části nervového systému a vysílají je do jiné části, kde mohou vyvolat nějakou činnost. Neurony mají rozličný tvar a velikost, ale základní složky mají všechny stejné. Z těla neuronu (viz. obrázek 2.) vystupuje množství jemných rozvětvených výběžků nazývaných dendrity a dlouhý, samostatný výběžek axon. Dendrity tvoří vstupy a axon tvoří výstupy. Axon jedné buňky se připojuje k ostatním buňkám pomocí dendritů přes synapse (je to spojení dvou buněk, z nichž alespoň jedna je nervová). Aktivní neuron vysílá z axonu elektronický signál, který prochází synapsí k ostatním neuronům, které se mohou stát aktivními [5], [24].

Neuronové síťe jsou dnes nedílnou součástí moderní informatiky. V poslední době dochází k velkému rozvoji a aplikaci neuronových sítí. Zabývají se jí nejen odborníci z oblasti informačních technologií, ale i z jiných oborů (finančnictví, jazykověda apod.). Neuronovou síť si můžeme z jistého pohledu představit jako funkci více proměnných, jejichž parametry jsou optimalizovány tak, aby poskytovaly hodnoty blízké



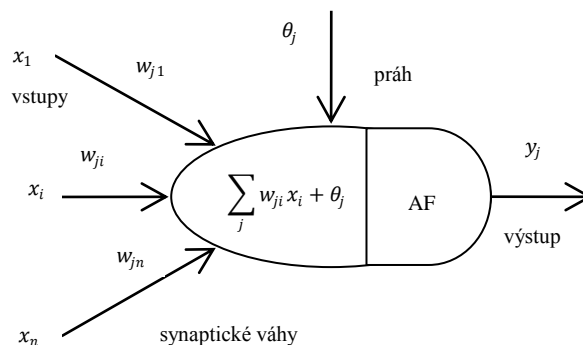
požadovaným. Proto může být neuronová síť využita pro klasifikaci a predikci [25], [26].



Obrázek 2: Neuron [12]

Jednotná definice pro UNS není známa, ale můžeme ji definovat jako orientovaný graf s dynamicky ohodnocenými hranami a vrcholy. Uzly grafu reprezentují neurony, které tvoří množinu  $\{X_1, X_2, \dots, X_n\}$  a spolu s množinou hran  $(i, j)$  grafu udávají topologie sítě. Každá hrana grafu je ohodnocena reálným číslem  $w_{ij}$  tzv. synaptickou váhou. Další definice je od americké společnosti DARPA,<sup>1</sup> ta říká, že: „Neuronová síť je systém sestávající z mnoha jednoduchých procesorů, pracujících paralelně, jejichž funkce je determinovaná strukturou sítě, intenzitou propojení a zpracováním ve výpočetních elementech nebo uzlech“ [15], [25], [26].

První formální model neuronu byl uveden v roce 1943 McCullochem a Pittsem. Jednalo se o binární prahový neuron. Model McCulloch-Pittsova neuron je nejrozšířenější model a označuje se jako formální neuron (viz. obrázek 3) [26].

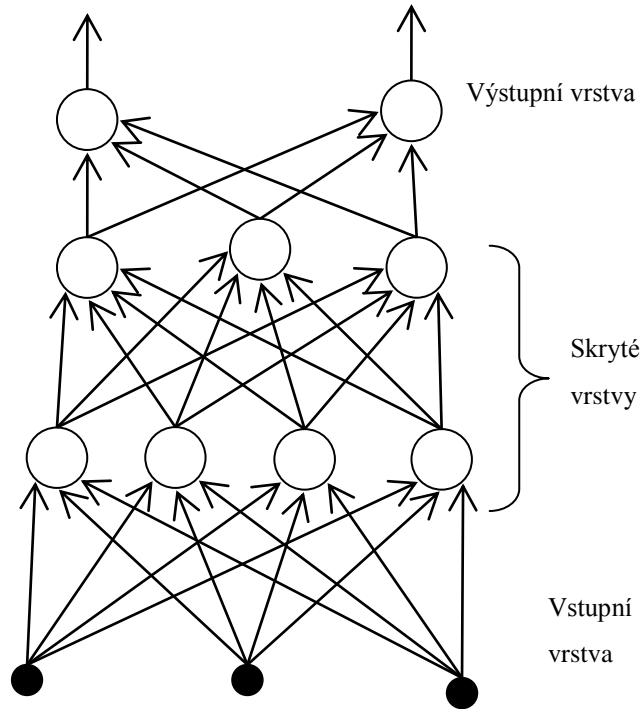


Obrázek 3: Model McCulloch-Pittsova neuronu [26]

<sup>1</sup> DARPA defence advance research project association

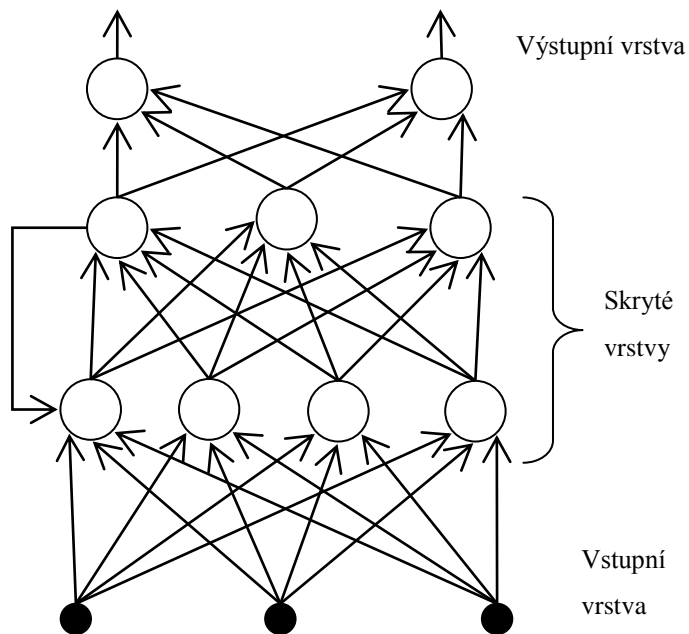
Podle propojení neuronu ve vrstvě rozlišujeme sítě [5], [24]:

- dopředním šířením informace (feedforward viz. obrázek 4);



Obrázek 4: Dopředná neuronová síť [25]

- rekurentní síť (recurrent network) neboli cyklické síť (viz. obrázek 5).



Obrázek 5: Rekurentní neuronová síť [25]

Dále můžeme UNS dělit podle počtu skrytých vrstev [5], [24]:

- s jednou skrytou vrstvou (viz. obrázek 6);

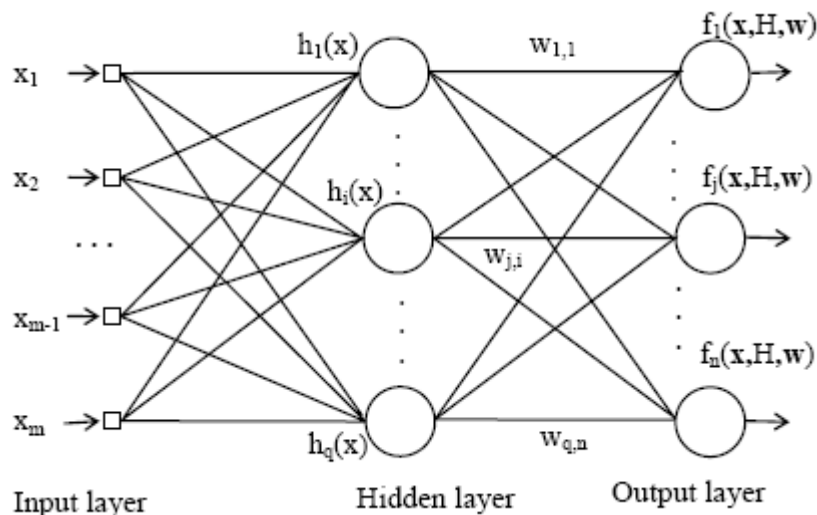
- vícevrstvé neuronové sítě (Multilayer Neural Networks) se skládají minimálně ze dvou vrstev (viz. obrázek 4).

S UNS je také spojen model perceptronu, který byl vymyšlen v 60. letech. Frank Rosenblatt ukázal, že McCullochové Pittsové modely s modifikovatelnými synaptickými váhami se dají natrénovat tak, aby uměly rozpoznat a klasifikovat objekty. Pojmem perceptron vyjadřuje vícevrstvou neuronovou síť s dopředními vazby, kde jsou všechny neurony jedné vrstvy propojeny se všemi neurony následující vrstvy. Do jednotlivého neuronu přichází impulzy jako reálná čísla, každý je vynásoben jemu příslušnou synaptickou váhou. Váha je také reálné číslo, které udává význam spoje. Impulzy vynásobené synaptickými váhami jsou sečteny, je odečten práh neuronu a na výsledek je aplikována aktivační funkce. Aktivační funkce může být skoková, spojitá lineární nebo spojitá nelineární [12], [15].

## 2.2 Topologie RBF neuronové sítě

RBF neuronová síť patří mezi dopředné neuronové sítě a odlišuje se o vícevrstvé neuronové sítě (MPL) tím, že má jen jednu skrytou vrstvu. Oblast použití je dost podobná MPL. Obě dvě jsou vhodné pro klasifikaci i regresi. Architektura RBF je jednoduchá, skládá se ze tří vrstev [1], [5], [8]:

- První vrstva je vstupní (input layer), kterou tvoří zdrojové uzly. Synapse mezi vstupní a skrytou vrstvou nejsou ohodnoceny váhami a vstup do RBF je nelineární.
- Druhá vrstva v síti je skrytá vrstva (hidden layer), která transformuje vstupní signál na hodnoty radiálních bazických funkcí. Ve skryté vrstvě mohou mít neurony různé aktivační funkce, přičemž nejčastější bývá využívaná Gaussova funkce.
- Výstupní vrstva (output layer) je odezva na vstupní signál. Je tvořena lineárními neurony, díky tomu je možno nahlížet na celou síť jako na určitý druh lineárního modelu.



Obrázek 6: RBF neuronová síť [22]

RBF můžeme definovat takto [22]:

$$f_j(x, H, w) = \sum_{i=1}^q w_{ij} \times h_i(x) \quad (1)$$

- kde:
- $f_j(x, H, w)$  je j-tý výstup;
  - $w_{ij}$  jsou váhy synapsí;
  - $H = (h_1, h_2, \dots, h_i, h_q)$  jsou aktivační funkce RBF;
  - vektor  $x = (x_1, x_2, \dots, x_k, x_m)$  je vstupní hodnota pro  $q$  aktivační funkce.

### 2.3 Aktivační funkce RBF neuronové sítě

Jak již bylo zmíněno, aktivační funkce využívá radiální bázovou funkci, která je zvláštní třídou matematické funkce. Její hlavní vlastností je monotónní stoupání nebo klesání s rostoucí vzdáleností od centra  $c_i$ . Je možné ji charakterizovat tímto obecným tvarem [22]:

$$h(x, C, R) = \sum_{i=1}^q \exp\left(-\frac{\|x - c_i\|^2}{r_i}\right) \quad (2)$$

- kde:
- $x = (x_1, x_2, \dots, x_k, \dots, x_m)$  představuje vstupní vektor;
  - $C = (c_1, c_2, \dots, c_i, \dots, c_q)$  jsou centrem aktivační funkce  $h_i(x)$  RBF neuronů;
  - $R = (r_1, r_2, \dots, r_i, \dots, r_q)$  jsou poloměry aktivační funkcí  $h_i(x)$ .

Neurony ve skryté vrstvě mohou jako aktivační funkci použít například Gaussovu funkci. Existuje mnoho radiálních bázových funkcí, ale pro RBF mají zvláštní význam následující funkce [5]:

- Multikvadratická

$$\varphi(r) = \sqrt{r^2 + c^2} \quad \text{kde } c > 0 \text{ a } r \in R \quad (3)$$

- Inverzní multikvadratická

$$\varphi(r) = \frac{1}{\sqrt{r^2+c^2}} \text{ kde } c > 0 \text{ a } r \in R \quad (4)$$

- Gaussova funkce

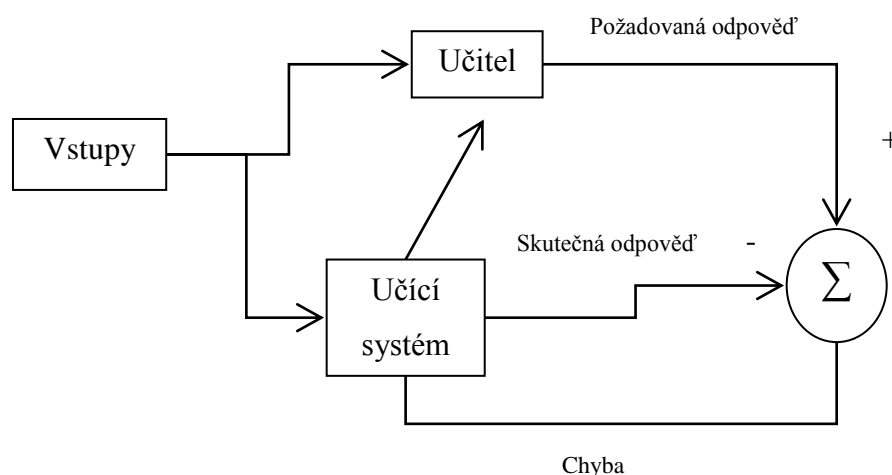
$$\varphi(r) = \exp\left(-\frac{r^2}{2\sigma^2}\right) \text{ kde } \sigma > 0 \text{ a } r \in R \quad (5)$$

## 2.4 Učení RBF neuronové sítě

Proces učení neuronové sítě je dynamický proces, při kterém dochází k nastavení parametrů. Cílem je dosažení shody mezi výstupy modelovaného systému a výstupy neuronové sítě. RBF se dokážou velmi rychle učit. Učící algoritmus je založen na učení s učitelem, a to i přes to, že stanovení středů RBF neuronů přímo nevyžaduje znalost požadovaných výstupních hodnot. Učení probíhá na základě trénovací množiny. Učení RBF sítě je rozděleno na fáze. V první fázi je určen počet center a pak nalezení nejvhodnější pozice RBF center. Další fází je určení poloměrů center, nastavení vah mezi skrytou a výstupní vrstvou a určení strmosti RBF [1], [6], [22].

### 2.4.1 Učení s učitelem

V praxi je to nejběžnější způsob učení, jestliže máme neuronovou síť se vstupními a výstupními neurony. Trénovací množina je tak tvořena dvojicí vektorů  $x$  a  $y$ . Trénovací množina obsahuje  $P$  příkladů, které se síť musí naučit. Učení, kde existuje nějaká vyšší autorita, která síti předkládá příklady dvojic vstupů a výstupů a pokud síť odpovídá špatně, informuje jí o chybě a modifikuje jí, se říká učení s učitelem (viz. obrázek 7) [5], [12].



Obrázek 7: Učení s učitelem [5]

## 2.4.2 Počet neuronů ve skryté vrstvě

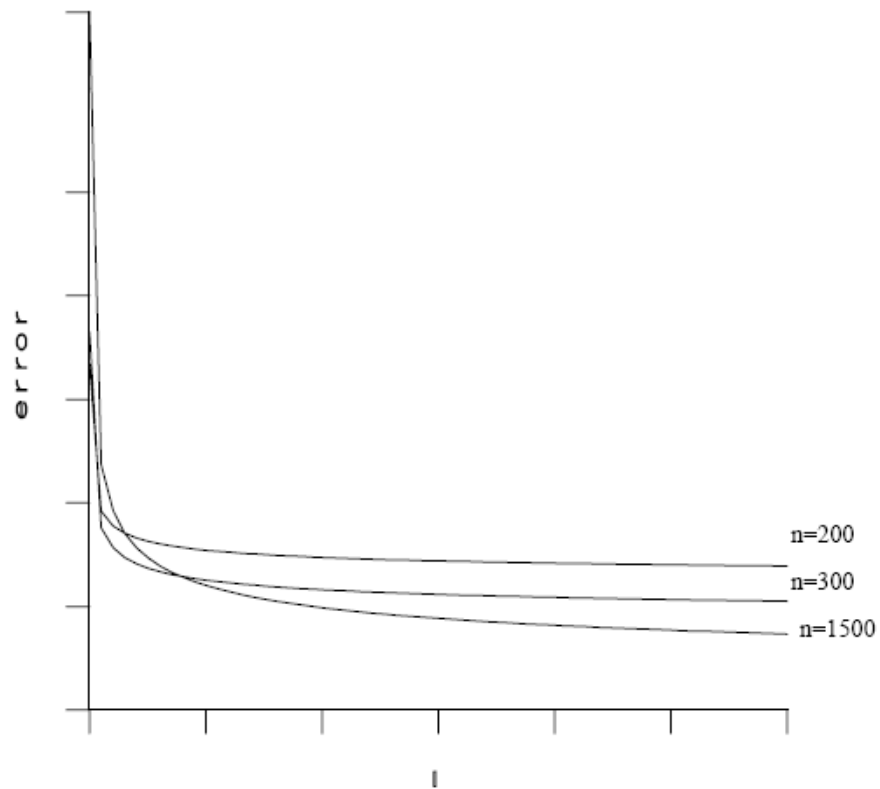
Počet neuronů ve skryté vrstvě má výrazný vliv na výsledné vlastnosti neuronové sítě, a to na rychlost učení a přesnost navržených modelů. První pokusy s neuronovou sítí RBF se zabývaly interpolací hodnot časových řad. Každé známé hodnotě byl přidělen jeden neuron ve skryté vrstvě. Pomocí tohoto postupu vznikly na výstupu neznámé hodnoty, které se nacházely mezi danými vstupy. Bohužel takto navržená neuronová síť je zcela nevhodná pro aproximaci různých funkcí, a to z důvodu, že vysoký počet neuronů vede k přeučení neuronové sítě. Přeučený model správně rozpozná pouze ty vzory, které se naučil a vzory, které se mírně liší, ty rozpoznat nedokáže. Počet skrytých vrstev se volí podle složitosti modelovaného systému. Problematikou stanovení počtu neuronů a jejich vlivu na přesnost se v devadesátých letech dvacátého století zabývali Feerico Girosi a Partha Niyogi. Výsledky jejich práce jsou podstatným přínosem, jelikož usnadňují návrh topologie neuronové sítě [4], [5], [6].

- Nechť je dána RBF neuronová síť používající Gaussovu funkci s  $m_0$  vstupními jednotkami a  $m_1$  skrytými jednotkami. Dále nechť  $f(x)$  představuje regresní funkci, jež patří do Sobolevova prostoru. Předpokládáme, že trénovací data jsou náhodný výběr z funkce  $f(x)$ , jež je funkcí regresní. Pro všechny parametry  $\delta \in (0,1)$  platí, že trénovací chyba  $O$  RBF neuronové sítě je ohraničena shora s pravděpodobností vyšší než  $1-\delta$  [5], [4], [9].

$$\|f_0 - \widehat{f_{m_1, N}}\|^2 \leq O\left(\frac{1}{m_1}\right) + O\left(\frac{m_0 \times m_1}{N} \times \log(m_1 N) + \frac{1}{N} \times \log\left(\frac{1}{\delta}\right)^{1/2}\right) \quad (6)$$

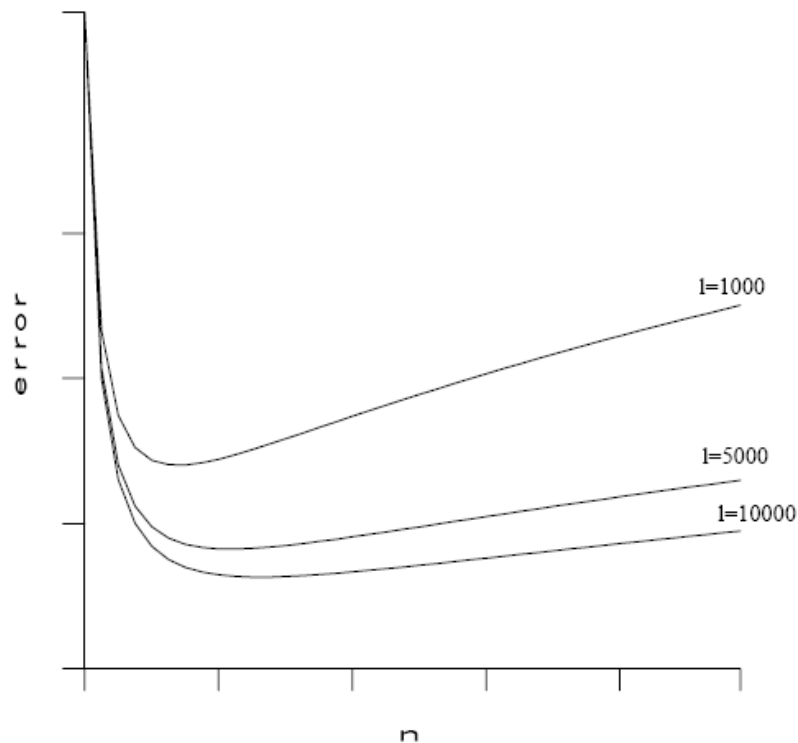
Na základě této věty můžeme dojít k těmto závěrům [5],[4]:

- Překrytím jednotlivých neuronů se stává neuronová síť RBF neefektivní pro řešení klasifikačních úloh.
- Počet neuronů ve skryté vrstvě má vliv na průběh trénovací chyby.
- S růstem počtu neuronu  $m_1$  (na obrázku 8 n představuje  $m_1$ ) ve skryté vrstvě celková chyba pro trénovací množinu dat konverguje k nule.



Obrázek 8: Vliv množství trénovacích dat a počet neuronů ve skryté vrstvě na trénovací chybu [4]

- Optimální množství neuronů ve skryté vrstvě pro trénovací množinu dat  $N$  je přibližně  $\sqrt[3]{N}$ . Udává také, že počet neuronů ve skryté vrstvě by měl být menší než dvakrát větší vstupní vrstva.
- Klesání testovací chyby je pouze v případě, že počet neuronů  $m_1$  roste výrazně pomaleji než množství dat  $N$  ve vstupní množině. V minimu je na křivce zevšeobecnovací chyba, ta se z množství dat použitých ve vstupní množině posouvá doprava.



Obrázek 9: Vliv počtu RBF neuronů na zevšeobecnující chybu [4]

### 2.4.3 Nalezení center RBF

Mezi nejdůležitější fáze učení je nalezení vhodných center RBF. Existuje řada způsobů, jak určit centra neuronové sítě typu RBF.

#### 2.4.3.1 Metoda náhodného výběru

Mezi metodami, jak určit centra, patří i metoda náhodného výběru z trénovacích dat. Tato metoda patří mezi nejjednodušší metody, jak určit centra. Tato metoda používá fixní sklon radiálně bazických funkcí. Jako radiálně bazickou funkci můžeme použít Gaussovu funkci, jejíž směrodatná odchylka je stanovena v závislosti na rozložení center. Konkrétně radiální bazická funkce se středem v  $t_i$  je definována jako [5].

$$G(\|x - t_i\|^2) = \exp\left(-\frac{m_1}{d_{\max}^2} \|x - t_i\|^2\right), \quad i = 1, 2, \dots, m \quad (7)$$

kde:  $-m_1$ : je počet center;

$-d_{\max}$ : je maximální vzdálenost mezi zvolenými centry.

Následujícím vztahem je zajištěno, že RBF nebude mít příliš plochý ani příliš strmý průběh [5].

$$\sigma = \frac{d_{\max}}{\sqrt{2m_1}} \quad (8)$$



Tato metoda je vhodná pro malé množství vstupních dat [5].

### 2.4.3.2 Segmentace

Z důvodu, že data, která vstupují do neuronové sítě typu RBF jsou předem známa, tak informace v nich obsažené mohou být využity k vyhledání vhodných míst jako center RBF oblastí. Existují dva možné postupy, jedním z nich je segmentace, který používá metodu učení bez učitele. Jeden z nejznámějších algoritmů shlukové analýzy je metoda K-means. Jedná se o nehierarchickou metodu. Postup se skládá s jednotlivých kroků [5], [7]:

- Na začátek se inicializují vektory  $u_j$ ,  $j = 1, \dots, k$ , použitím nějaké vhodně zvolené heuristiky nebo náhodně zvolenými hodnotami.
- Všechna data se přiřadí k určeným vektorům podle minima Euklidovské vzdálenosti.

$$d(X, Y) = \sqrt{\sum_{i=1}^p (x_i - y_i)^2} \quad (9)$$

- Vypočítají se nové hodnoty vektorů jako střední hodnoty dat  $x_i$ , která se přiřadila k určitému vektoru.
- Tyto kroky opakujeme do té doby, dokud dochází k přesunu.

Klasifikace dat je další metoda, kterou můžeme použít, jedná se tak o metodu učení s učitelem.

## 2.5 Všeobecný klasifikační problém

Zavedená všeobecná formulace klasifikačního problému je zobrazena pomocí funkce zobrazení. Funkce zobrazení je definována nad dvěma množinami A a B. Necht'  $F(x)$  je funkce definovaná nad množinou A, která přiřadí každému elementu  $x \in A$  obraz funkčních hodnot z množiny B,  $\hat{x} = F(x) \in B$ ,

$$F: A \rightarrow B \quad (10)$$

Necht'  $G(x, w)$  je funkce, jejíž argumenty jsou z konečné podmnožiny  $A_{\text{tren}} = \{x_1, x_2, \dots, x_r\} \subset A$  a  $w$  je parametr zobrazení  $G$ , potom  $\hat{x} = G(x, w) \in B_{\text{tren}} \subset B$ ,

$$G(w): A_{\text{tren}} \rightarrow B_{\text{tren}} \quad (11)$$

Je možné říci, že zobrazení  $G(w)$  je restrikcí zobrazení  $F(x)$  nad množinou  $A_{\text{tren}} \subset A$ . Komplement  $A_{\text{tren}}$  vzhledem k množině A je označen  $A_{\text{test}}$  (nazývána testovací

množinou),  $A_{\text{test}}=A/A_{\text{tren}}$ . Je předpoklad, že pro každé  $x_i \in A_{\text{tren}}$ , je znám požadovaný obraz, funkčních hodnot  $\hat{x}_i$

$$\frac{x_1}{\hat{x}_1}, \frac{x_2}{\hat{x}_2}, \dots, \frac{x_r}{\hat{x}_r}. \quad (12)$$

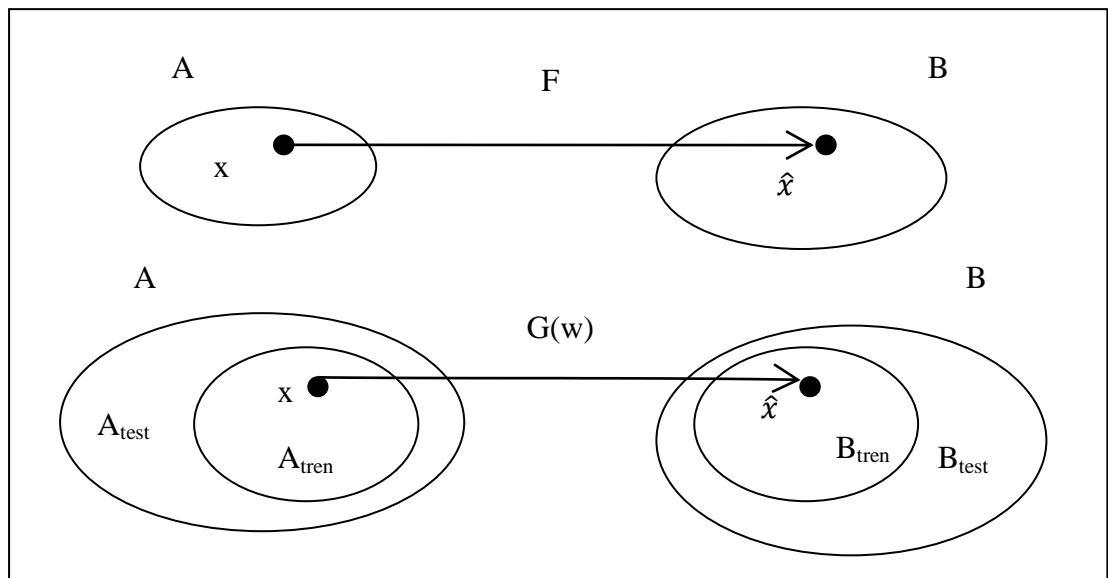
Požadované funkční hodnoty  $\hat{x}_i$  jsou poté obrazy funkce  $F$  [13].

$$\hat{x}_i = F(x_i), (i=1,2,\dots,r). \quad (13)$$

Cílem je nalezení takového parametru  $w$  funkce  $G(x,w)$ , aby funkční hodnoty argumentů z trénovací množiny  $A_{\text{tren}}$  byly co nejbližší obrazům funkce  $F(x)$ . Cílem je tedy nalezení minima funkce [13].

$$E(w) = \frac{1}{2} \sum_{i=1}^r (G(x_i, w) - \hat{x}_i)^2 \quad (14)$$

Tato funkce vyjadřuje sumu kvadratických odchylek funkce  $G(x,w)$  od hodnot  $\hat{x}_i$  trénovací množiny [13].



Obrázek 10: Schematické zobrazení  $F: A \rightarrow B$  [13]

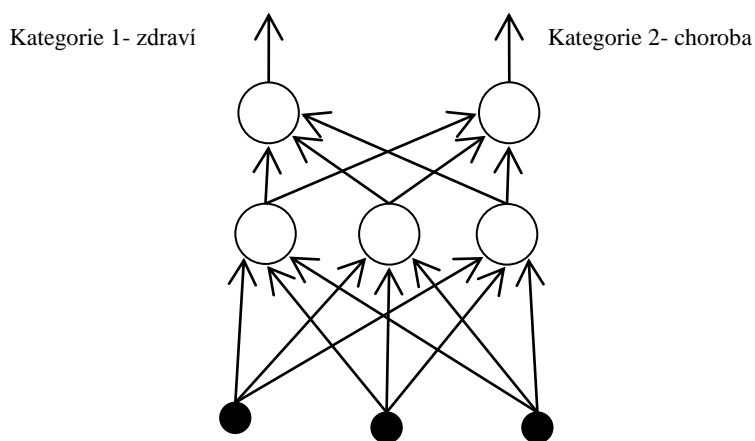
Všeobecný klasifikační problém je možné vysvětlit tak, že pro natrénování neuronové sítě je třeba mít sadu vstupů a zároveň odpovídající sadu už známých výstupu. Trénovací množina je použita při trénování neuronové sítě. Testovací množina je použita pro hodnocení, jak úspěšně se podařilo neuronovou sít' natrénovat. Při trénování je snaha, aby suma čtverců odchylek výstupů z neuronové sítě byla co nejmenší od předem zadaných hodnot. Tato suma je označena jako účelová funkce. Její minimalizace je dosažena změnou nějakého parametru či parametrů neuronové sítě [14].

## 2.6 Klasifikace

Klasifikace je sdružení dat do tříd (počet tříd může být předem znám). Klasifikovat je možno data [26]:

- s učitelem: jedná se o klasifikaci, která je založena na porovnávání neznámých vzorů  $x$  se všemi známými vzory  $x_k$  na základě kriterií určujících stupeň podobnosti či vzdálenosti v nějaké metrice. Při této klasifikaci je znám počet tříd, do kterých budeme klasifikovat. Jednou z nejdůležitějších klasifikačních metod je metoda K-nejbližšího souseda. Další metodou je klasifikace vzorů podle úhlu.
- bez učitele: není znám počet tříd, do kterých budou vzory klasifikovány. Vzájemné vztahy mezi vzory jsou určovány pomocí metod shlukování. Dělí se na jednoduché metody a hierarchické metody. Ke klasifikaci jevů dochází na základě podobnosti. Tyto metody se uplatňují zejména při automatizaci analýzy mnohorozměrných dat.

Použití UNS pro klasifikaci má uplatnění v mnoha oblastech, jeden z příkladů, který se uvádí, se týká medicíny. Skupina doktorů vzala množství záznamů z EKG. Každý z lékařů měl určit, zda snímek patří člověku trpící srdeční chorobou nebo ne. Následně tyto záznamy dávali do neuronové sítě na vstup a učili ji správně odpovídat. Když byla neuronová síť na trénovací množině záznamů naučena, začali jí předkládat záznamy, které předtím neviděla. Ukázalo se, že UNS pozná anomálii tak dobře, jako by rozhodovala skupina lékařů. Síť tak prováděla klasifikaci záznamů na dvě skupiny (viz. obrázek 11) [12].



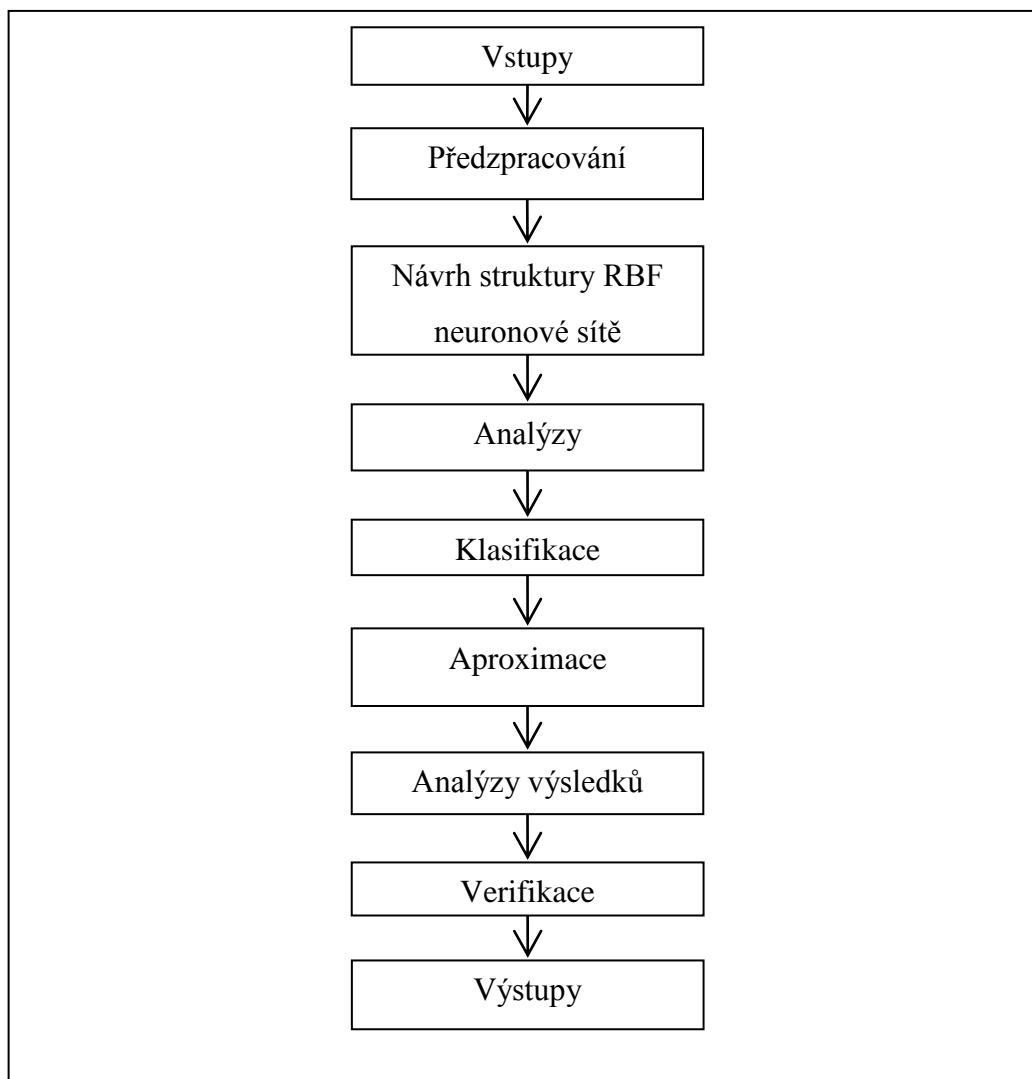
Obrázek 11: RBF síť pro klasifikaci [8]

Klasifikovat pomocí sítě můžeme i do více skupin.

### 3 Návrh modelu udržitelného rozvoje

Model je zjednodušeným zobrazením reality. Následující obrázek znázorňuje návrh modelu pro udržitelný rozvoj. První část této kapitoly je zaměřena na analýzu parametrů. Další fáze se zabývá předzpracováním dat v prostředí Excelu. Data musela být nejdříve ošetřena a pak byla rozdělena na trénovací a testovací množinu. Následné zpracování dat v prostředí SPSS Clementine.

Pro modelování udržitelného rozvoje byl navrhnout model, který je znázorněn na obrázku 12. Model se skládá z jednotlivých činností: předzpracování, návrhu struktury RBF, analýzy, klasifikace, aproximace a analýzy výsledků. Na závěr byla provedena verifikace.



Obrázek 12: Model udržitelného rozvoje [zdroj: vlastní]

## 3.1 Analýza parametrů

Vstupními daty (parametry) je popsán udržitelný rozvoj v obcích ve vybrané lokalitě. Vybranou lokalitou je Pardubický kraj, kde je evidováno 451 obcí, z toho 15 obcí s rozšířenou působností a 26 obcí s pověřeným obecním úřadem. Indikátory (parametry) byly navrženy expertem přes danou problematiku. Data byla rozdělena podle základních tří pilířů udržitelného rozvoje [10].

### 3.1.1 Ekonomický pilíř

Ekonomický pilíř je především soustředěn na posilování konkurenceschopnosti ekonomiky za předpokladu respektování limitu existujících environmentálních zdrojů. Ekonomické parametry (EKO) reprezentující udržitelný rozvoj jsou znázorněny v tabulce 1[11].

Tabulka 1: Ekonomické parametry udržitelného rozvoje

Ekonomické parametry	
Podíl ekonomicky aktivních	EKO1
Koncentrace ekonomiky	EKO2
Průměrná mzda/průměr kraje	EKO3
Dluhová služba	EKO4
Dluh na obyvatele	EKO5
Podíl opakujících se příjmů a běžných výdajů	EKO6
Podíl vlastních příjmů	EKO7
Podíl kapitálových výdajů	EKO8
Likvidní majetek na obyvatele	EKO9
Přijaté dotace na obyvatele	EKO10
Počet podnikatelských subjektů/ tis. obyvatel	EKO11

Navrženými parametry z ekonomické oblasti jsou [1], [3], [10], [28]:

- **Podíl ekonomicky aktivních:** jde o poměr mezi počtem lidí, kteří jsou ekonomicky aktivní a celkovým počtem obyvatel obce. Mezi ekonomicky aktivní patří všechny zaměstnané osoby (tj. všechny osoby 15leté a starší, které v rozhodný okamžik sčítání byly v placeném zaměstnání jako zaměstnanci, zaměstnavatelé, podnikatelé a osoby samostatně výdělečně činné), osoby ve vazbě a ve výkonu trestu odnětí svobody, pracující studenti a učni, ženy na mateřské dovolené (tj. pouze ženy na mateřské dovolené v trvání maximálně 28, resp. 37 týdnů, které jsou v pracovním poměru) a

nezaměstnaní (tj. osoby 15leté a starší, které v rozhodný okamžik sčítání byly bez práce, hledaly aktivně práci a byly připraveny k okamžitému nástupu do práce).

- **Koncentrace ekonomiky:** tento indikátor nám popisuje koncentraci ekonomiky u obcí. Malá hodnota indikátoru znamená dlouhodobou pružnost místní ekonomiky a ochranu proti úpadku jednoho sektoru.
- **Průměrná mzda/průměr kraje:** jde o poměr mezi průměrnou mzdou v obci a průměrnou mzdou v kraji. Velikost mezd má vliv na poptávku.
- **Dluhová služba:** označuje roční platby úroku. Jestliže je hodnota indikátoru nad 0,15, lze to považovat za hrozící signál. Ukazatelé dluhové služby slouží k monitorování hospodaření obcí. I když v roce 2008 bylo zrušeno usnesení vlády o Regulaci zadluženosti obcí a krajů pomocí ukazatele dluhové služby, tak některá ministerstva nadále vyžadují předkládání výpočtu ukazatele dluhové služby. Ukazatel dluhové služby je dán tímto vztahem:

$$\text{Ukazatel dluhové služby} = \text{dluhová služba} / \text{dluhová základna} \quad (15)$$

$$\text{Dluhová služba} = \text{úroky} + \text{splátky jistin a dluhopisů} + \text{splátky leasing} \quad (16)$$

$$\text{Dluhová základna} = \text{daňové příjmy} + \text{nedaňové příjmy} + \text{přijaté dotace} \quad (17)$$

- **Dluh na obyvatele:** jedná se poměr mezi celkovým dluhem v korunách a počtem obyvatel.
- **Podíl opakujících se příjmů a běžných výdajů:** tento indikátor vypovídá o kvalitě rozpočtového hospodaření. Když je hodnota větší než 1, rozpočtový přebytek je běžný.
- **Podíl vlastních příjmů:** jedná se o podíl vlastních příjmů na celkových příjmech. Vlastní příjmy tvoří příjmy daňové, nedaňové a kapitálové. Z 90% celkových vlastních příjmů tvoří vlastní příjmy daňové výnosy. Do vlastních příjmů obce nezahrnují dotace. Vyšší je podíl vlastních příjmů na celkové příjmy. Znamená to vyšší fiskální autonomii obce. Čím vyšší je fiskální autonomie obce, tím mají obce menší potřebu volit dluh jako nástroj financování. Mezi nedaňové příjmy patří např. platby za odebrané množství podzemní.
- **Podíl kapitálových výdajů:** ukazatel charakterizuje objem investovaných prostředků. Kapitálové výdaje lze rozdělit do dvou skupin, a to na kapitálové výdaje hrazené z účelových dotací získaných ze státního rozpočtu a výdaje

hrazené z vlastních zdrojů. Zvláštní část potom tvoří kapitálové výdaje financované spoluúčastí evropských fondů.

- **Likvidní majetek na obyvatele:** každá z obcí je vlastníkem nějakého majetku, ten je často použit jako zástava bankovních úvěrů. Likvidním majetkem se pak rozumí majetek, na který si obec může vzít úvěr.
- **Přijaté dotace na obyvatele:** jedná se o podíl celkové přijaté dotace obce a počtu obyvatel obce.
- **Počet podnikatelských subjektů na tisíc obyvatel:** jde o podíl celkového počtu registrovaných podnikatelských subjektů v obci na 1000 obyvatel.

### 3.1.2 Environmentální pilíř

Environmentální pilíř se zaměřuje na ochranu životního prostředí, resp. jeho zlepšování a zabránění vyčerpání omezených přírodních zdrojů. Indikátory v charakterizující environmentální oblasti jsou znázorněny v tabulce 2 [11].

Tabulka 2: Environmentální parametry (EVN)

Environmentální parametr	
Kanalizace s ČOV	EVN1
Kanalizace bez ČOV	EVN2
Podíl orné půdy ze zemědělské půdy (%)	EVN3
Podíl zastavěných a ostatních ploch z celkové výměry (%)	EVN4
Podíl vodních ploch z celkové výměry (%)	EVN5
Koeficient ekologické stability	EVN6
Zahrady na obyvatele (ha/obyv.)	EVN7
Ovocné sady na obyvatele (ha/1 obyv.)	EVN8
Trvalé travní porosty (ha/obyv.)	EVN9
Lesní půda na obyvatele (ha/obyv.)	EVN10
Zemědělská půda na obyvatele (ha/obyv.)	EVN11

Navrženými parametry z environmentální oblasti jsou [1], [3], [28]:

- **Kanalizace s ČOV:** tento indikátor nabývá logických hodnot 0 nebo 1 podle toho, zda obec má kanalizaci s čističkou odpadních vod. Hodnota 1 představuje kladnou odpověď, 0 pak zápornou odpověď.
- **Kanalizace bez ČOV:** tento indikátor nabývá také logických hodnot 0 nebo 1 podle toho, zda obec má kanalizaci bez čističky odpadních vod. Hodnota 1 představuje kladnou odpověď, 0 pak zápornou odpověď.

- **Podíl orné půdy ze zemědělské půdy v %:** zachycuje podíl orné půdy, která patří mezi nestabilní krajinnotvorné prvky, na celkové výměře zemědělské půdy.
- **Podíl zastavěných a ostatních ploch z celkové výměry (%)** je dán podílem součtů výměr zastavěných ploch a ostatních ploch v daném území a celkové výměry území.
- **Koeficient ekologické stability:** je podíl ekologicky příznivých ploch a ploch, které zatěžují životní prostředí. Počítá se jako poměr druhů pozemků (chmelnice + vinice + zahrady + ovocné sady + trvalé travní porosty + pastviny + lesní půda + vodní plochy) / (orná půda + zastavěné plochy + ostatní plochy)).
- **Podíl vodních ploch z celkové výměry (%):** je dán podílem výměry vodních ploch v daném území a celkové výměry území.
- **Zahrady na obyvatele (ha/obyvatel):** jde o podíl rozlohy zahrad v dané obci k celkovému počtu obyvatel v dané obci. Za zahradu považujeme pozemek, který je zpravidla oplocený a na kterém se převážně pěstuje zelenina, květiny a jiné zahradní plodiny, zpravidla pro vlastní spotřebu. Zahrady jsou také souvislé pozemky osázené ovocnými stromy nebo keři až do výměry 0,25ha. Zahrada může být součástí obytné nebo hospodářské budovy.
- **Ovocné sady na obyvatele (ha/obyvatel.):** jde o podíl rozlohy ovocných sadů v dané obci k celkovému počtu obyvatel v dané obci. Minimální výměra ovocných sadů 0,15 ha je stanovena ve Vyhlášce Českého statistického úřadu č. 126/2001, kterou se stanoví kritéria pro zápis do zemědělského registru.
- **Trvalé travní porosty (ha/ obyvatel):** jde o podíl rozlohy trvalých travnatých porostů v dané obci k celkovému počtu obyvatel v dané obci. Kdysi byl tento parametr označován jako ukazatel louky.
- **Lesní půda na obyvatele (ha/ obyvatel):** tento parametr udává podíl rozlohy lesní půdy k celkovému počtu obyvatel obce.
- **Zemědělská půda na obyvatele (ha/ obyvatel.)** tento parametr udává podíl rozlohy zemědělské půdy k celkovému počtu obyvatel obce.

### 3.1.3 Sociální pilíř

Sociální pilíř je oproti předešlým pilířům více zaměřený na přítomnost. Jeho vazba na budoucí generaci je znatelná pouze v případě problematiky stárnutí populace. Sociální pilíř by měl sledovat především zajištění kvalitního života z hlediska absence hmotné



nouze. V některých případech do problematiky vstupuje i uspokojování kulturních potřeb. Indikátory reprezentující sociální oblast jsou uvedeny v tabulce 3 [11].

Tabulka 3: Sociální parametry (SOC)

Sociální parametr	
Vyjíždějící do zaměstnání/počet obyvatel	SOC1
Vyjíždějící do škol/počet obyvatel.	SOC2
Dojíždějící do zaměstnání/počet obyvatel	SOC3
Dojíždějící do škol/počet obyvatel.	SOC4
Počet trvale obydlených bytů/počet obyvatel	SOC5
Počet bytů k rekreaci/počet obyvatel	SOC6
Podíl obyvatel s pitnou vodou	SOC7
Podíl obyvatel s plynem	SOC8
Podíl obyvatel napojených na kanalizaci	SOC9
Počet obyvatel	SOC10
Míra nezaměstnanosti	SOC11
Podíl VŠ obyvatel	SOC12
Stanice vlaku	SOC13
Přirozený přírůstek	SOC14
Saldo migrace	SOC15
Podíl obyvatel ve věku 0 -14 na celkovém počtu obyvatel (%)	SOC16
Podíl obyvatel ve věku 65 a více let na celkovém počtu obyvatel (%)	SOC17
Počet uchazečů o zaměstnání/ počet obyvatel	SOC18
Počet dokončených bytů/počet obyvatel	SOC19
Počet lázeňských léčeben	SOC20

Navrženými parametry ze sociální oblasti jsou [3], [28]:

- **Vyjíždějící do zaměstnání/počet obyvatel:** tento parametr zobrazuje podíl počtu vyjíždějících lidí, což jsou lidé, kteří mají v obci trvalé bydliště, ale do zaměstnání dojíždějí do jiné obce, k celkovému počtu obyvatel obce.
- **Vyjíždějící do škol/počet obyvatel:** tento parametr zobrazuje podíl počtu vyjíždějících studentů, kteří mají v obci trvalé bydliště, ale do školy dojíždějí do jiné obce, k celkovému počtu obyvatel obce.
- **Dojíždějící do zaměstnání/počet obyvatel:** tento parametr udává podíl počtu dojíždějících zaměstnanců, což jsou lidé, kteří nemají v obci trvalé bydliště, ale do obce dojíždějí za prací, k celkovému počtu obyvatel obce.
- **Dojíždějící do škol/počet obyvatel:** tento parametr udává podíl počtu dojíždějících studentů, což jsou studenti, kteří nemají v obci trvalé bydliště, ale do obce dojíždějí do školy, k celkovému počtu obyvatel obce.

- **Počet trvale obydlených bytů/počet obyvatel:** tento parametr zobrazuje podíl počtu trvale obydlených bytů k celkovému počtu obyvatel obce. Trvale obydlený byt je ten, kde je nahlášena minimálně jedna osoba s trvalým bydlištěm a zároveň tento byt splňuje podmínky stavebního úřadu.
- **Počet bytů k rekreaci/počet obyvatel:** tento parametr zobrazuje podíl počtu bytů sloužících k rekreaci k celkovému počtu obyvatel obce. Jedná se převážně o chaty, chalupy apod.
- **Podíl obyvatel s pitnou vodou:** tento parametr zobrazuje podíl počtu obyvatel s trvalým bydlištěm v obci, které mají byt zásobovaný pitnou vodou, k celkovému počtu obyvatel obce.
- **Podíl obyvatel s plynem:** tento parametr zobrazuje podíl počtu obyvatel s trvalým bydlištěm v obci, které mají přípojku na plyn, k celkovému počtu obyvatel obce.
- **Podíl obyvatel napojených na kanalizaci:** tento parametr zobrazuje podíl počtu obyvatel s trvalým bydlištěm v obci, které jsou napojeny na místní kanalizaci, k celkovému počtu obyvatel obce.
- **Počet obyvatel:** představuje celkový počet obyvatel za určité období.
- **Míra nezaměstnanosti:** slouží pro hodnocení situace na trhu práce a pro hodnocení vývoje v oblasti rozvoje lidských zdrojů. Míra nezaměstnanosti vyjadřuje podíl počtu nezaměstnaných za dané období v dané obci a pracovní síly (ekonomický aktivní). Pracovní síla představuje součet počtu nezaměstnaných a zaměstnaných obyvatel za dané období na daném území. Nezaměstnanými mohou být osoby starší 15 let, které v daném období nebyly zaměstnány a aktivně si hledaly práci, do které byly schopny nastoupit nejpozději do 14dnů.
- **Podíl VŠ obyvatel:** je poměr vysokoškolsky vzdělaných obyvatel obce k celkovému počtu obyvatel obce.
- **Stanice vlaku:** tento parametr nabývá buď hodnoty 1, nebo 0. Když je parametr ohodnocen hodnotou 1, tak v obci je zřízená stanice vlaku, v opačném případě obec nemá stanici.
- **Přirozený přírůstek:** rozdíl mezi počtem živě narozených a zemřelých během roku.

- **Saldo migrace:** vyjadřuje rozdíl mezi počtem obyvatel, kteří se do obce přistěhovali a počtem obyvatel, kteří z obce emigrovali.
- **Podíl obyvatel ve věku 0-14 na celkovém počtu obyvatel (%):** je poměr počtu obyvatel obce, jejichž věk je v rozmezí 0 až 14 k celkovému počtu obyvatel vyjádřený v procentech.
- **Podíl obyvatel ve věku 65 a více let na celkovém počtu obyvatel (%):** je poměr počtu obyvatel obce, jejichž věk je od 65let a více, k celkovému počtu obyvatel vyjádřený v procentech.
- **Počet uchazečů o zaměstnání/ počet obyvatel:** jedná se o poměr počtu uchazečů o zaměstnání v obci, kteří jsou evidováni na úřadu práce, k celkovému počtu obyvatel.
- **Počet dokončených bytů/počet obyvatel:** tento parametr zobrazuje podíl počtu dokončených bytů k celkovému počtu obyvatel obce. Dokončeným bytem se označuje byt, který prošel nějakou rekonstrukcí a byl zkolaudován.
- **Počet lázeňských léčeben:** udává počet lázeňských léčeben v daných obcích

## 3.2 Předzpracování dat

Před začátkem jakékoliv operace s daty musí být data nejprve ošetřena. Nejdříve bude provedena standardizace a pak normalizace [16], [24]:

- **Standardizace:** k tomu, aby mohly být rozděleny jednotlivé objekty do tříd, musí být všechny parametry souměřitelné. Z důvodu, že jednotlivé indikátory jsou v různých jednotkách, musí být provedena standardizace, která potlačí citlivost na měřítka a dá parametrům stejnou váhu. Je vycházeno z toho, že matice má řádky a sloupce, pak standardizace bude provedena podle následujícího vztahu.

$$Z_{ij} = \frac{x_{ij} - \bar{x}_j}{S_j} \quad (18)$$

- kde :
- $i = 1, 2, \dots, m$ ;
  - $j = 1, 2, \dots, n$ ;
  - $\bar{x}_j$  = střední hodnota (průměr);
  - $S_j$  = směrodatná odchylka;
  - $x_{ij}$  = prvek.

Kontrola správně provedené standardizace poslouží k vypočítání střední hodnoty, které se musí u všech parametrů rovnat. Standardizované znaky jsou pak bezrozměrné.

- **Normalizace:** dále je provedena normalizace na potlačení velkých rozdílů mezi prvky. Nejdříve musí být ze standardizované matice vypočítána norma, ze které jsou pak vypočítány hodnoty normalizované matice.

$$a_i = \sqrt{\left[ \sum_{j=1}^s (x_{ij})^2 \right]} \quad (19)$$

kde:  $a_i$  = norma.

Normalizovaná matice je vypočítána podle následujícího vzorce.

$$X_{ij} = \frac{x_{ij}}{a_i} \quad (20)$$

Dále byla množina objektů rozdělena na trénovací a testovací množinu. Rozklad množiny objektů na trénovací a testovací patří mezi základní problémy při aplikacích neuronových sítí [13].

$$A = A_{\text{tren}} \cup A_{\text{test}} \quad (21)$$

Trénovací množina by měla obsahovat objekty z množiny A, které dobře reprezentují ostatní objekty zahrnuté v testovací množině. Jeden ze způsobů, jak rozdělit množinu na trénovací a testovací, je pomocí klastrovací metody (shlukové metody). Pomocí shlukové metody je možno rozložit množinu A na disjunktí podmnožiny neboli klastry, které obsahují “podobné” objekty [13].

$$A = C_1 \cup C_2 \cup \dots \cup C_n \quad (22)$$

kde i-ty klastr  $C_i$  obsahuje  $n_i$  objektů z A

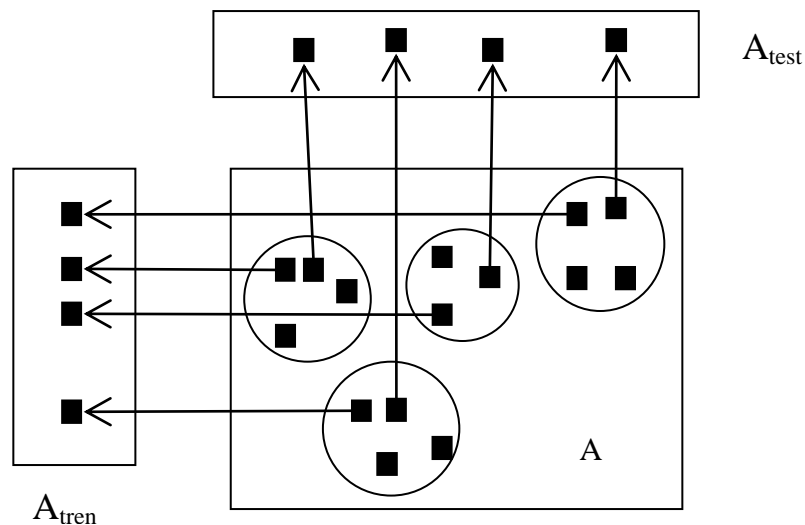
$$C_i = \{o_1^i, o_2^i, \dots, o_{n_i}^i\} \subset A \quad (23)$$

Objekt  $o_1^i \in C_i$  je objekt z i-tého klastru  $C_i$ , který leží nejbližší k jeho centru. Tento objekt bude následně sloužit jako “reprezentant” objektu z klastru  $C_i$ . Potom trénovací a testovací množina je určena objekty [13].

$$A_{\text{tren}} = \{o_1^1, o_2^1, \dots, o_{n_1}^1\} \quad (24)$$

$$A_{\text{test}} = (C_1 - \{o_1^{(1)}\}) \cup (C_2 - \{o_1^{(2)}\}) \cup \dots \cup (C_n - \{o_1^{(n)}\}) \quad (25)$$

To znamená, že trénovací množina je složená ze všech klastrů a testovací množina obsahuje zbylé objekty (viz obrázek 1). Počet objektů v trénovací množině je totožný s počtem klastrů,  $|A_{\text{tren}}|=p$  a  $|A_{\text{test}}|=|A|-p$  [13].



Obrázek 13: Rozklad množiny objektů na trénovacích a testovacích množiny [13]

Mezi další metody, jak rozdělit data na trénovací a testovací patří [9], [8]:

- výběr na základě oblastí, které jsou určeny Kohonenovou samoorganizující sítí;
- rozdělení náhodným výběrem;
- výběr každého  $n$ -tého prvku.

Podle trénovacích dat se nastavují u neuronových sítí váhy synapsí. Podle testovacích dat se pozná, jak se neuronová síť naučila. Validační data slouží ke kontrole po učení [21].

V prostředí SPSS Clementine umožňuje rozdělit data na trénovací a testovací uzel Partitoin. Generátoru umožňuje rozdělení dat náhodným výběrem. V příloze 1 jsou porovnání rozdělení trénovacích a testovacích dat jednotlivých modelů podle tříd. Data jsou rozdělena do 7. shluků (clasty) pomocí Kohonenovy samoorganizující sítě. Tyto shluky pomáhají při učení neuronové sítě, jedná se tak o učení s učitelem.

### 3.3 Návrh struktury RBF neuronové sítě

U neuronové sítě typu RBF se používá metoda učení s učitelem. Data jsou klasifikována do předem daného nastaveného počtu tříd. Učení s učitelem umožňuje porovnávat momentální skutečnou odezvu sítě s odezvou žádoucí.

U navržených modelů (ekonomický, environmentální a sociální) byly obce v Pardubickém kraji klasifikovány do 7. tříd, kde [22]:

- V první třídě se shluky nejlepší obce, např. u ekonomického modelu jsou to obce s velmi příznivými ekonomickými podmínkami a nízkou zadlužeností.

- Do druhé třídy patří obce, které jsou velmi dobré. V ekonomickém modelu jsou to obce, které umí dostát svým závazkům.
- Ve třetí třídě jsou dobré obce. To jsou obce, které jsou schopny z hlediska ekonomického modelu dostát svým finančním závazkům.
- Do třetí třídy patří stabilní obce. V ekonomickém modelu jsou to obce se střední zadlužeností.
- V páté třídě jsou obce, které už nejsou stabilní, z ekonomického hlediska jsou to obce, které plní své závazky jen za příznivých ekonomických podmínek.
- V šesté třídě jsou obce z hlediska ekonomiky s vysokou zadlužeností.
- V sedmé třídě jsou ty nejhorší obce, které z ekonomického hlediska jsou neschopné dostát svým finančním závazkům.

Vstupem do neuronové sítě byly u každého modelu jeho parametry, které jsou popsány v kapitole 3.1. Výstupem se také stal vektor správných klasifikačních tříd (pomocí něho poznáme, jak se neuronová síť naučila), ten nabývá jen hodnot 0 nebo 1 (dummy proměnná).

U RBF neuronové sítě je nastaven počet neuronů ve skryté vrstvě, počet učicích cyklů tak, aby nedošlo k přeučení, což by vedlo k růstu chyb. O nastavení jednotlivých parametrů neuronové sítě RBF se zabývá následující kapitola.

## 4 Analýza výsledků navržené struktury RBF

Tato kapitola je rozdělena na čtyři části. První tři části pojednávají o nastavení jednotlivých modelů, které byly vytvořeny (ekonomické, environmentální a sociální).

U všech modelů jsou popsány přesnosti klasifikace, jak na trénovacích datech na kterých probíhá proces učení, tak na testovacích datech. Na nich poznáme, jak je systém naučený. Podle přesnosti na testovacích datech probíhal výběr optimálního návrhu neuronové sítě RBF. Závěrem je celkové zhodnocení všech tří modelů. Realizace návrhu neuronové sítě byla provedena v prostředí Clementine a následné výsledky byly zpracované v prostředí MS-Excel.

Clementine nabízí nastavení neuronové sítě prostřednictvím uzlu Neural Net. V tomto uzlu je možné si vybrat metodu např. RBFN, která představuje RBF neuronovou síť. Po vybrání metody RBFN uzel umožňuje nastavit tyto parametry:

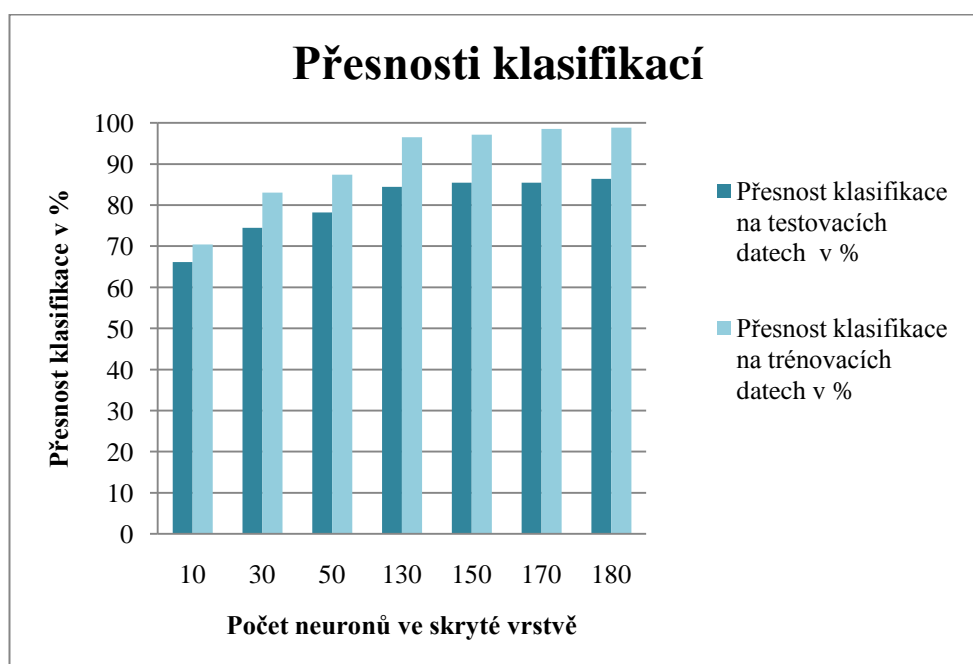
- počet neuronů ve skryté vrstvě;
- parametry alpha;
- překrytí neuronové sítě;
- persistence;
- počet cyklů.

### 4.1 Ekonomický model

Pro ekonomický model byla data rozdělena takto: 75% trénovací a 25% testovací. První parametr, který byl zkoumán se týkal nastavení neuronů ve skryté vrstvě. Přehled nastavení neuronů ve skryté vrstvě je popsán v následující tabulce 4. Tabulka popisuje, jak se se změnou počtu neuronů měnily i přesnosti klasifikací neuronové sítě. Výsledky jsou pro lepší přehled vyobrazeny na grafu 1. Z grafu vyplývá, že optimální počet neuronů ve skryté vrstvě je 180. Při tomto nastavení činily přesnosti klasifikací na trénovacích datech 98,851% a na testovacích datech 86,408%.

Tabulka 4: Počet neuronů ve skryté vrstvě pro ekonomický model [zdroj: vlastní]

Počet neuronů ve skryté vrstvě	Persistence	$\alpha$	Překrytí RBF vrstvy	Počet cyklů	Přesnost klasifikace na trénovacích datech v %	Přesnost klasifikace na testovacích datech v %
10	30	0,9	1	500	70,440	66,165
30	30	0,9	1	500	83,019	74,436
50	30	0,9	1	500	87,421	78,195
130	30	0,9	1	500	96,542	84,466
150	30	0,9	1	500	97,118	85,437
170	30	0,9	1	500	98,559	85,437
180	30	0,9	1	500	98,851	86,408



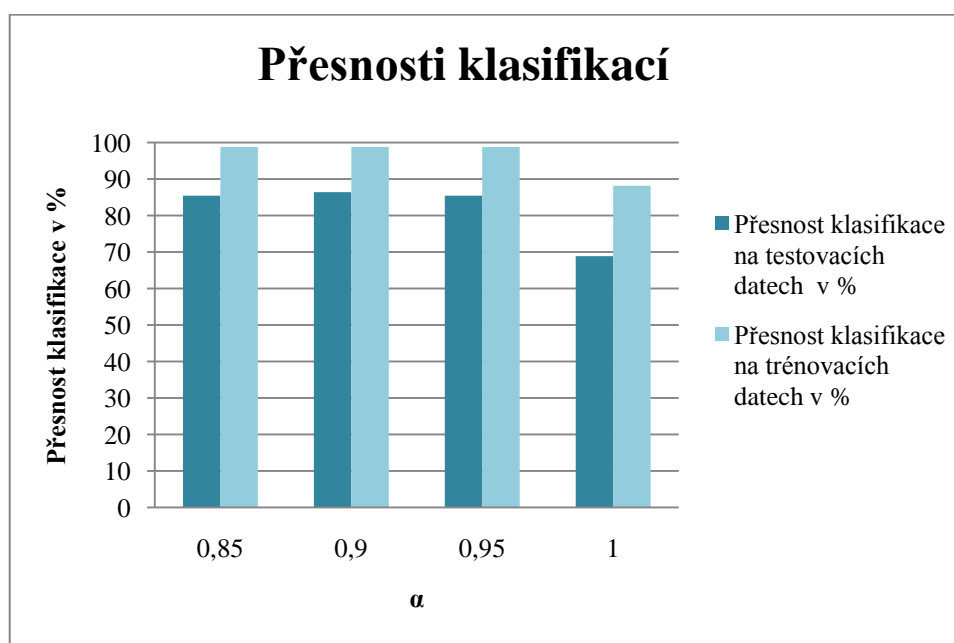
Graf 1: Přesnosti klasifikací při změně počtu neuronů ve skryté vrstvě pro ekonomický model [zdroj: vlastní]

Parametr  $\alpha$  byl další parametr, který byl zkoumán. Jedná se o parametr momentu. Tento parametr může být nastaven v intervalu  $\langle 0,1 \rangle$ . Přehled nastavení parametrů  $\alpha$  je v následující tabulce 5. Tabulka popisuje, jak se se změnou parametru  $\alpha$  měnily i přesnosti klasifikací neuronové sítě. Výsledky jsou znázorněny na grafu 2, ze kterého vyplývá, že optimální parametr  $\alpha$  je při hodnotě 0,9. Přesnosti klasifikací na trénovacích datech činila 98,851% a na testovacích datech 86,408%.



Tabulka 5: Parametr  $\alpha$  pro ekonomický model [zdroj: vlastní]

Počet neuronů ve skryté vrstvě	Persistence	$\alpha$	Překrytí RBF vrstvy	Počet cyklů	Přesnost klasifikace na trénovacích datech v %	Přesnost klasifikace na testovacích datech v %
180	30	0,85	1	500	98,851	85,437
180	30	0,9	1	500	98,851	86,408
180	30	0,95	1	500	98,851	85,437
180	30	1	1	500	88,218	68,932

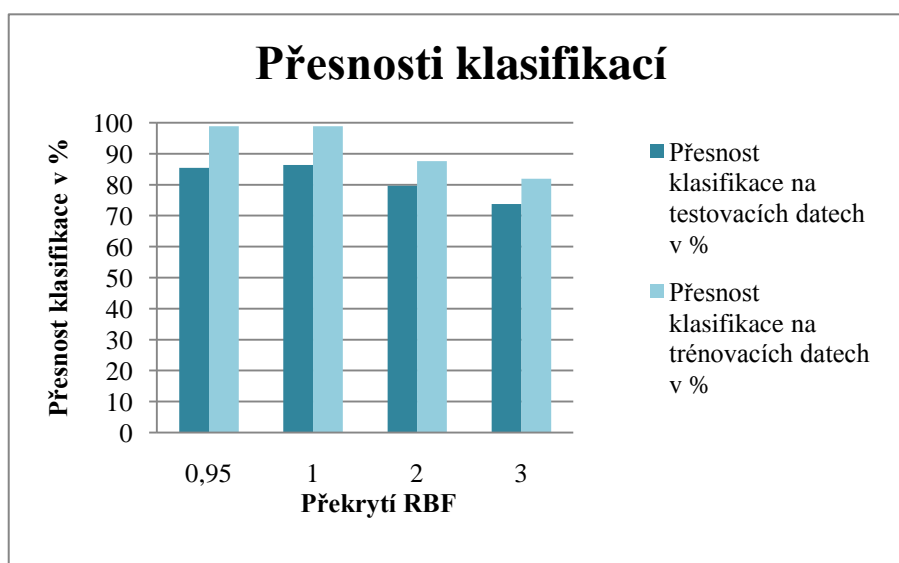


Graf 2: Přesnosti klasifikací při změně parametru  $\alpha$  pro ekonomický model [zdroj: vlastní]

Následně bylo zkoumáno, jak přesnosti klasifikací ovlivní parametr překrytí RBF. Přesnosti klasifikací jsou znázorněny v tabulce 6. Pro lepší přehled jsou výsledky znázorněny na grafu 3, ze kterého vyplývá, že optimální překrytí je při hodnotě 1. Při tomto nastavení činily přesnosti klasifikací na trénovacích datech 98,851% a na testovacích datech 86,408% .

Tabulka 6: Překrytí RBF pro ekonomický model [zdroj: vlastní]

Počet neuronů ve skryté vrstvě	Persistence	$\alpha$	Překrytí RBF vrstvy	Počet cyklů	Přesnost klasifikace na trénovacích datech v %	Přesnost klasifikace na testovacích datech v %
180	30	0,9	0,95	500	98,851	85,437
180	30	0,9	1	500	98,851	86,408
180	30	0,9	2	500	87,644	79,612
180	30	0,9	3	500	81,897	73,786

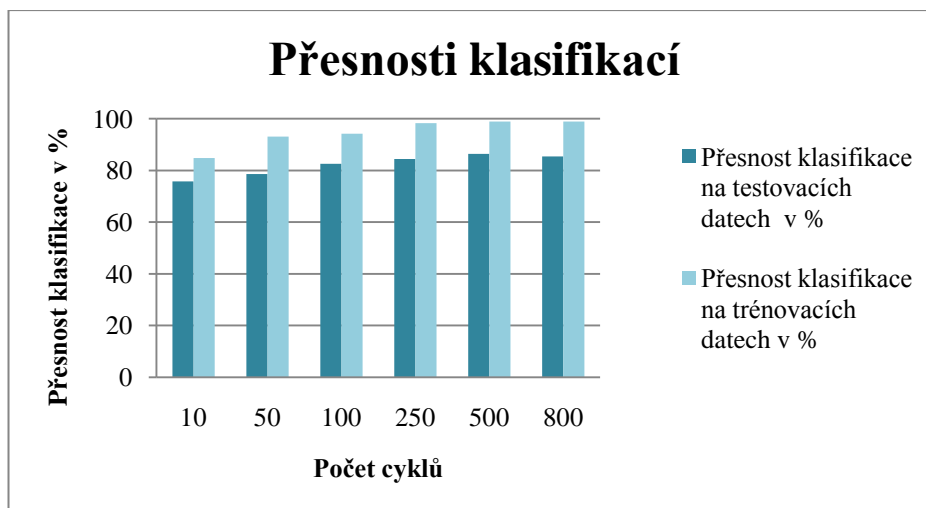


**Graf 3: Přesnosti klasifikací při změně překrytí RBF pro ekonomický model [zdroj: vlastní]**

Poslední parametr byl počet cyklů. Přehled nastavení počtu cyklů je v následující tabulce 7. Tabulka popisuje, jak se se změnou počtu cyklů měnily i přesnosti klasifikací neuronové sítě. Výsledky jsou znázorněny na grafu 4, ze kterého vyplývá, že optimální počet cyklů je 500, při nichž jsou přesnosti klasifikací na trénovacích datech 98,851% a na testovacích datech 86,408% .

**Tabulka 7: Počet cyklů pro ekonomický model [zdroj: vlastní]**

Počet neuronů ve skryté vrstvě	Persistence	$\alpha$	Překrytí RBF vrstvy	Počet cyklů	Přesnost klasifikace na trénovacích datech v %	Přesnost klasifikace na testovacích datech v %
180	30	0,9	1	10	84,770	75,728
180	30	0,9	1	50	93,103	78,641
180	30	0,9	1	100	94,253	82,524
180	30	0,9	1	250	98,276	84,466
180	30	0,9	1	500	98,851	86,408
180	30	0,9	1	800	98,851	85,437



Graf 4: Přesnosti klasifikací při změně počtů cyklů pro ekonomický model [zdroj: vlastní]

Při změnách parametru persistence nedošlo k žádným změnám přesnosti klasifikací (viz tabulka 8), proto byl tento parametr vyřazen z analýzy.

Tabulka 8: Persistence pro ekonomický model [zdroj: vlastní]

Počet neuronů ve skryté vrstvě	Persistence	$\alpha$	Překrytí RBF vrstvy	Počet cyklů	Přesnost klasifikace na trénovacích datech v %	Přesnost klasifikace na testovacích datech v %
180	30	0,9	1	500	98,851	86,408
180	130	0,9	1	500	98,851	86,408

Prostřednictvím analýzy, jež byla provedena na ekonomických datech, bylo získáno optimální nastavení neuronové sítě RBF, které je znázorněno v tabulce 9. Přesnosti klasifikací činily u trénovacích dat 98,851% , u testovacích dat 86,408%.

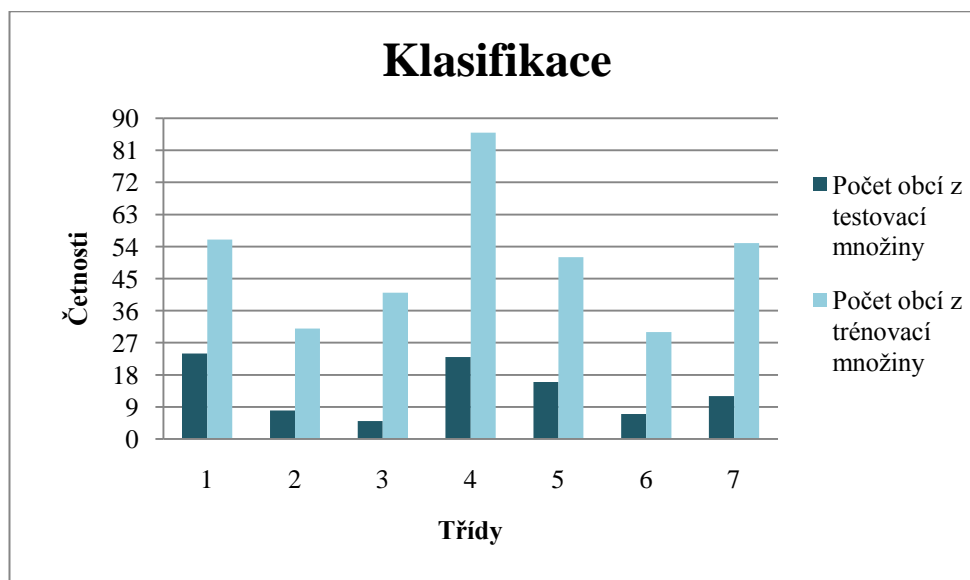
Tabulka 9: Optimální nastavení pro ekonomický model [zdroj: vlastní]

Počet neuronů ve skryté vrstvě	Persistence	$\alpha$	Překrytí RBF vrstvy	Počet cyklů	Přesnost klasifikace na trénovacích datech v %	Přesnost klasifikace na testovacích datech v %
180	30	0,9	1	500	98,851	86,408

V následující tabulce 10 je přehled výsledných klasifikací obcí. Obce z trénovací a testovací množiny jsou klasifikovány do jednotlivých tříd. Pro lepší přehled jsou klasifikace znázorněné pomocí grafu 5. Z grafu vyplývá, že nejvíc obcí z trénovací množiny se nachází ve 4. třídě, tam připadlo 86 obcí. Z testovací množiny připadlo nejvíc obcí do 1. třídy, tam připadlo 24 obcí. Z trénovací množiny připadlo nejméně obcí do 6. třídy, tam připadlo 30 obcí. Z testovací množiny připadlo nejméně obcí do 3. třídy pouze 5 obcí.

Tabulka 10: Klasifikace ekonomického modelu [zdroj: vlastní]

Třídy	1	2	3	4	5	6	7
Počet obcí z testovací množiny	24	8	5	23	16	7	12
Počet obcí z trénovací množiny	56	31	41	86	51	30	55



Graf 5: Klasifikace ekonomického modelu [zdroj: vlastní]

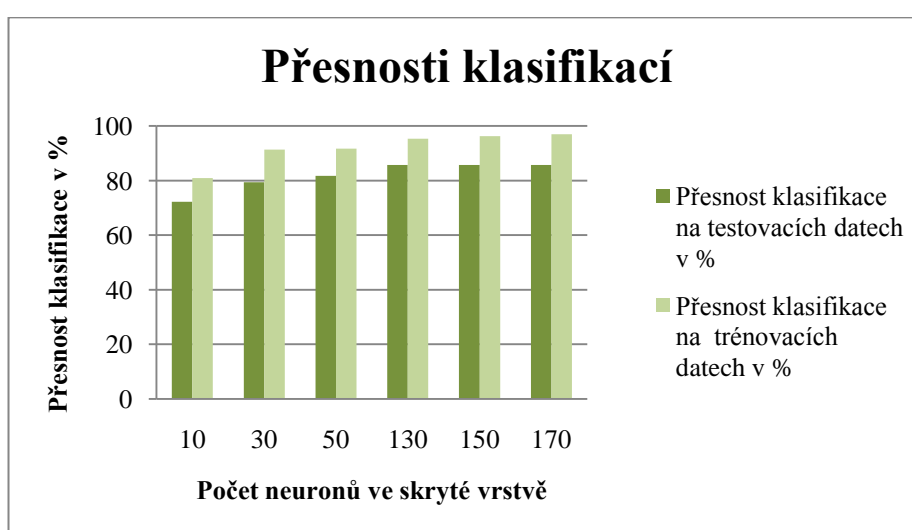
## 4.2 Environmentální model

Pro environmentální model byla data rozdělena takto: 75% trénovací a 25% testovací. Stejný postup analýzy, který byl použit u ekonomického modelu, byl použit také u environmentálního modelu.

Jako první parametr byl zkoumán počet neuronů ve skryté vrstvě (viz tabulka11). Tabulka popisuje, jak se se změnou počtu neuronů měnily i přesnosti klasifikací neuronové sítě. Přesnost klasifikace na testovacích datech se ustálila, když počet neuronů ve skryté vrstvě dosáhl na hodnotu 130, ale přesnost klasifikace na trénovacích datech se nadále zvětšovala. Z tohoto důvodu bylo optimální nastavení neuronů ve skryté vrstvě 170. Výsledky jsou vyobrazeny na grafu 6. Přesnosti klasifikací byly na trénovacích datech 96,914% a na testovacích datech 85,714%.

Tabulka 11: Počet neuronů ve skryté vrstvě pro environmentální model [zdroj: vlastní]

Počet neuronů ve skryté vrstvě	Persistence	$\alpha$	Překrytí RBF vrstvy	Počet cyklů	Přesnost klasifikace na trénovacích datech v %	Přesnost klasifikace na testovacích datech v %
10	30	0,9	2	500	80,864	72,222
30	30	0,9	2	500	91,358	79,365
50	30	0,9	2	500	91,667	81,746
130	30	0,9	2	500	95,370	85,714
150	30	0,9	2	500	96,296	85,714
170	30	0,9	2	500	96,914	85,714

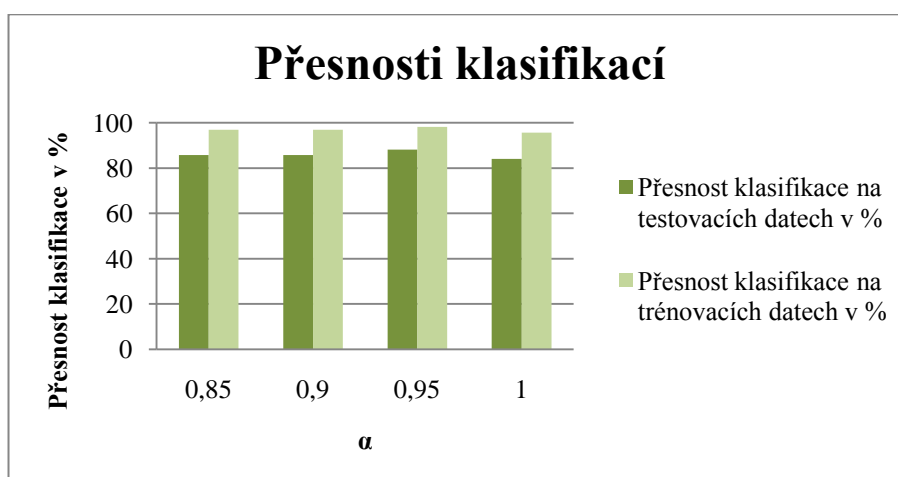


Graf 6: Přesnosti klasifikací při změně počtu neuronů ve skryté vrstvě pro environmentální model [zdroj: vlastní]

Přehled nastavení parametrů  $\alpha$  je v následující tabulce 12, která popisuje, jak se se změnou parametru  $\alpha$  měnily i přesnosti klasifikací neuronové sítě. Výsledky jsou znázorněny na grafu 7, ze kterého vyplývá, že optimální parametr  $\alpha$  je při hodnotě 0,95. Přesnosti klasifikací činily na trénovacích datech 98,148% a na testovacích datech 88,095%.

Tabulka 12: Parametr  $\alpha$  pro environmentální model [zdroj: vlastní]

Počet neuronů ve skryté vrstvě	Persistence	$\alpha$	Překrytí RBF vrstvy	Počet cyklů	Přesnost klasifikace na trénovacích datech v %	Přesnost klasifikace na testovacích datech v %
170	30	0,85	2	500	96,914	85,714
170	30	0,9	2	500	96,914	85,714
170	30	0,95	2	500	98,148	88,095
170	30	1	2	500	95,679	84,127

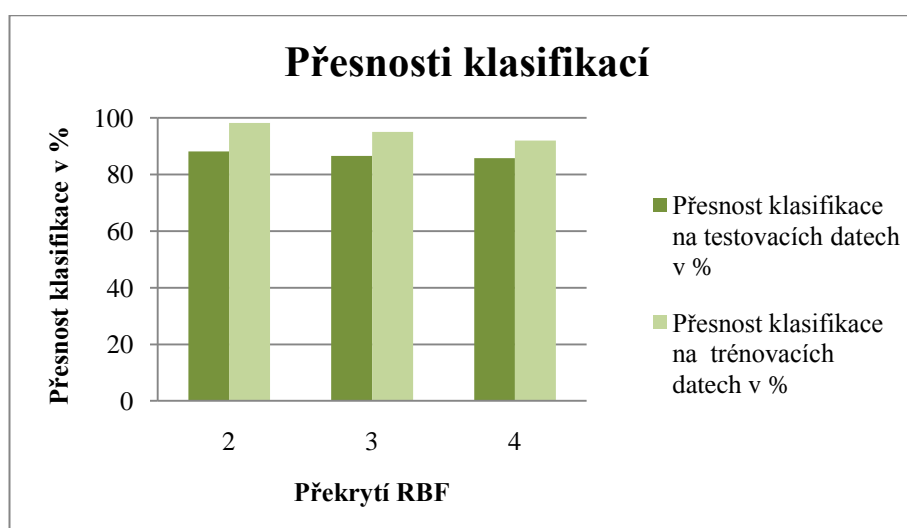


**Graf 7: Přesnosti klasifikací při změně parametru  $\alpha$  pro environmentální model [zdroj: vlastní]**

Následně bylo zkoumáno, jak jsou přesnosti klasifikací ovlivněny parametrem překrytí RBF. Přesnosti klasifikací jsou znázorněny v tabulce 13. Pro lepší přehled slouží graf 8. Optimální překrytí je při hodnotě 2, kdy přesnosti klasifikací činily na trénovacích datech 98,148% a na testovacích datech 88,095%.

**Tabulka 13: Překrytí RBF pro environmentální model [zdroj: vlastní]**

Počet neuronů ve skryté vrstvě	Persistence	$\alpha$	Překrytí RBF vrstvy	Počet cyklů	Přesnost klasifikace na trénovacích datech v %	Přesnost klasifikace na testovacích datech v %
170	30	0,95	2	500	98,148	88,095
170	30	0,95	3	500	95,062	86,508
170	30	0,95	4	500	91,975	85,714



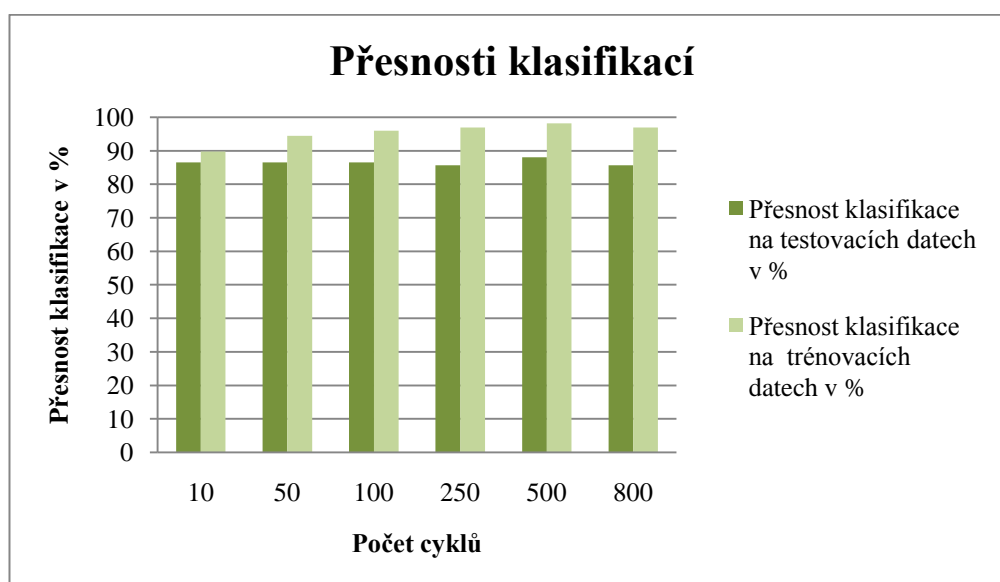
**Graf 8: Přesnosti klasifikací při změně překrytí RBF pro environmentální model [zdroj: vlastní]**

Poslední parametr byl počet cyklů. Přehled nastavení počtů cyklů je v následující tabulce 14. Tabulka popisuje, jak se se změnou počtu cyklů měnily i přesnosti

klasifikace neuronové sítě. Výsledky jsou znázorněny na grafu 9, ze kterého vyplývá, že optimální počet cyklu je 500. Přesnosti klasifikací činily na trénovacích datech 98,148% a na testovacích datech 88,095%.

Tabulka 14: Počet cyklů pro environmentální model [zdroj: vlastní]

Počet neuronů ve skryté vrstvě	Persistence	$\alpha$	Překrytí RBF vrstvy	Počet cyklů	Přesnost klasifikace na trénovacích datech v %	Přesnost klasifikace na testovacích datech v %
170	30	0,95	2	10	89,815	86,508
170	30	0,95	2	50	94,444	86,508
170	30	0,95	2	100	95,988	86,508
170	30	0,95	2	250	96,914	85,714
170	30	0,95	2	500	98,148	88,095
170	30	0,95	2	800	96,914	85,714



Graf 9: Přesnosti klasifikací při změně počtů cyklů pro environmentální model [zdroj: vlastní]

Pomocí analýzy, která byla provedena na environmentálních datech, bylo nakonec získáno optimální nastavení (viz. tabulka 15). Přesnosti klasifikací činily na trénovacích datech 98,148% a na testovacích datech 88,095%.

Tabulka 15: Optimální nastavení pro environmentální model [zdroj: vlastní]

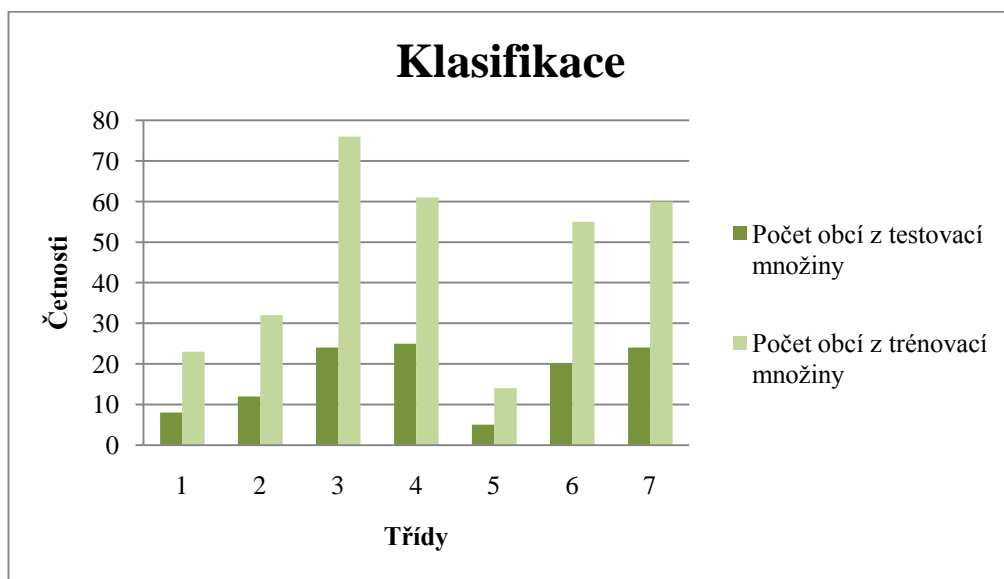
Počet neuronů ve skryté vrstvě	Persistence	$\alpha$	Překrytí RBF vrstvy	Počet cyklů	Přesnost klasifikace na trénovacích datech v %	Přesnost klasifikace na testovacích datech v %
170	30	0,95	2	500	98,148	88,095

V následující tabulce 16 je přehled výsledné klasifikací obcí. Obce z trénovací a testovací množiny jsou klasifikovány do jednotlivých tříd. Pro lepší přehled jsou klasifikace znázorněny pomocí grafu 9. Obce z trénovací množiny připadly nejvíc do 3.

třídy, kam připadlo 76 obcí. U testovací množiny byly obce nejčastěji klasifikovány do 4. třídy, kde se nachází 25 obcí. Nejméně obcí bylo přiřazeno do 5. třídy. Z trénovací množiny bylo přiřazeno do této třídy 14 obcí a z testovací množiny 5 obcí.

Tabulka 16: Klasifikace environmentálního modelu [zdroj: vlastní]

Třídy	1	2	3	4	5	6	7
Počet obcí z testovací množiny	8	12	24	25	5	20	24
Počet obcí z trénovací množiny	23	32	76	61	14	55	60



Graf 10: Klasifikace environmentálního modelu [zdroj: vlastní]

### 4.3 Sociální model

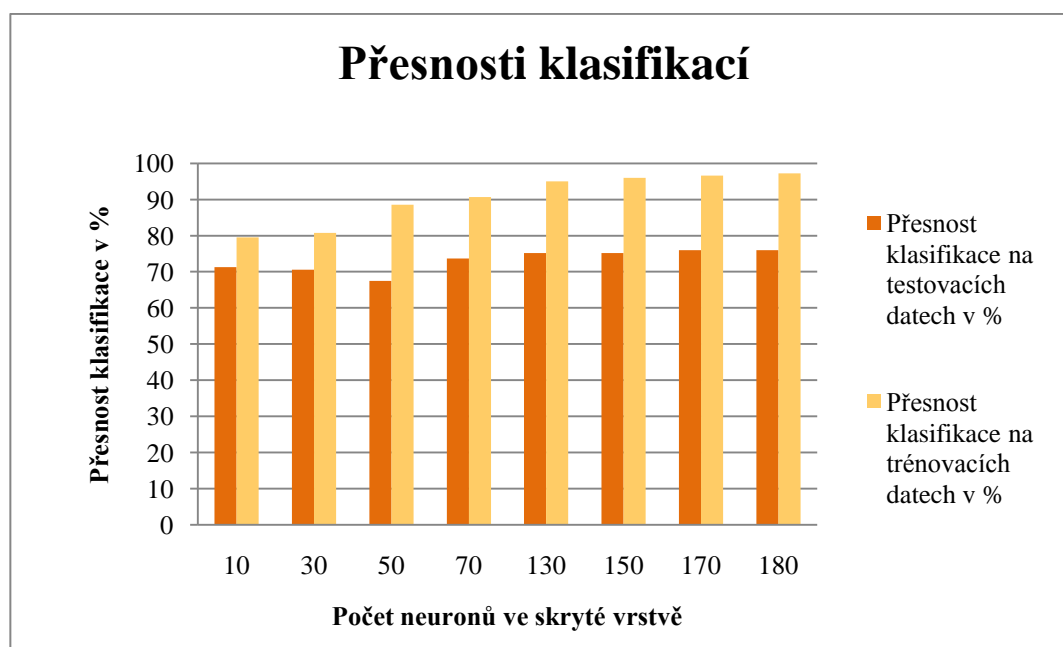
Pro ekonomický model byla data rozdělena takto: 72% trénovací a 28% testovací. Stejný postup analýzy, který byl použit u ekonomického modelu, byl použit také u sociálního modelu.

Jako první parametr, byl zkoumán počet neuronů ve skryté vrstvě (viz. tabulka 17). Tabulka popisuje, jak se se změnou počtu neuronů měnily i přesnosti klasifikací neuronové sítě. Výsledky jsou pak vyobrazeny na grafu 11. Z grafu vyplývá, že optimální počet neuronů ve skryté vrstvě je 180. Při této hodnotě přesnosti klasifikací byly na trénovacích datech 97,205% a na testovacích datech 75,969%.



Tabulka 17: Počet neuronů ve skryté vrstvě pro sociální model [zdroj: vlastní]

Počet neuronů ve skryté vrstvě	Persistence	$\alpha$	Překrytí RBF vrstvy	Počet cyklů	Přesnost klasifikace na trénovacích datech v %	Přesnost klasifikace na testovacích datech v %
10	30	0,9	1	500	79,503	71,318
30	30	0,9	1	500	80,745	70,543
50	30	0,9	1	500	88,509	67,442
70	30	0,9	1	500	90,683	73,643
130	30	0,9	1	500	95,031	75,194
150	30	0,9	1	500	95,963	75,194
170	30	0,9	1	500	96,584	75,969
180	30	0,9	1	500	97,205	75,969

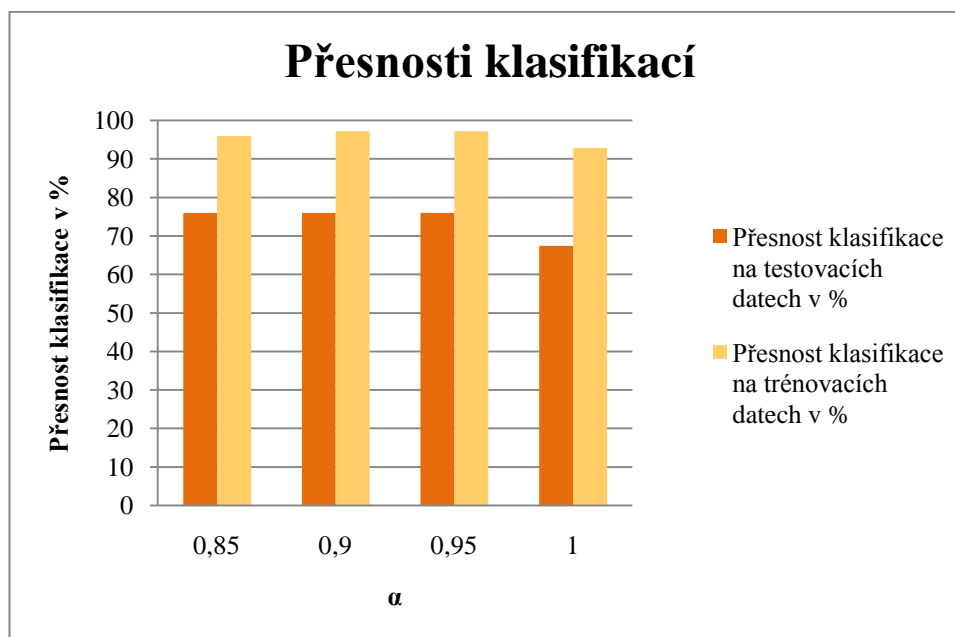


Graf 11: Přesnosti klasifikací při změně počtu neuronů ve skryté vrstvě pro sociální model [zdroj: vlastní]

Dále byl zkoumán parametr  $\alpha$ , což je parametr momentu. Přehled nastavení parametrů  $\alpha$  je v následující tabulce 18. Tabulka popisuje, jak se se změnou parametru  $\alpha$  měnily i přesnosti klasifikací neuronové sítě. Výsledky jsou znázorněny na grafu 12, ze kterého vyplývá, že optimální parametr  $\alpha$  je při hodnotě 0,9. Přesnosti klasifikací byly na trénovacích datech 97,205% a na testovacích datech 75,969%.

Tabulka 18: Parametr  $\alpha$  pro sociální model [zdroj: vlastní]

Počet neuronů ve skryté vrstvě	Persistence	$\alpha$	Překrytí RBF vrstvy	Počet cyklů	Přesnost klasifikace na trénovacích datech v %	Přesnost klasifikace na testovacích datech v %
180	30	0,85	1	500	95,963	75,969
180	30	0,9	1	500	97,205	75,969
180	30	0,95	1	500	97,205	75,969
180	30	1	1	500	92,857	67,442

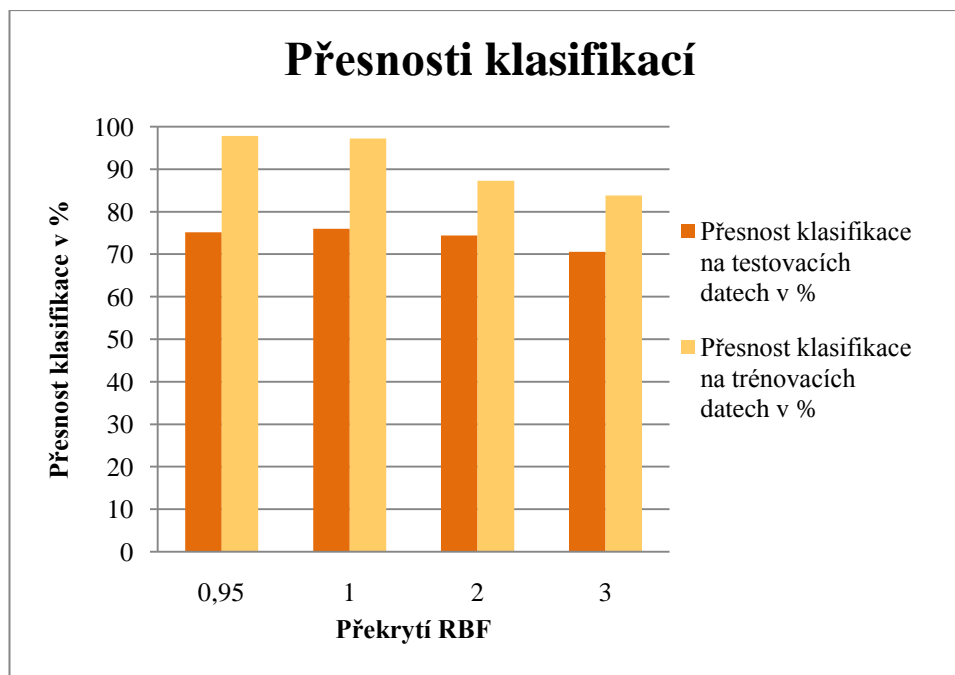


Graf 12: Přesnosti klasifikací při změně parametru  $\alpha$  pro sociální model [zdroj: vlastní]

Následně bylo zkoumáno, jak přesnost klasifikace ovlivní parametr překrytí RBF. Přesnosti klasifikací jsou znázorněny v tabulce 19, pro lepší přehled slouží graf 3. Přestože u trénovacích dat byla největší přesnost klasifikace při hodnotě 0,95, byly preferované přesnosti klasifikace testovacích dat. Na nich poznáme, jak se systém naučil. Z tohoto důvodu bylo optimální překrytí při hodnotě 1, kdy přesnosti klasifikací byly na trénovacích datech 97,205% a na testovacích datech 75,969%.

Tabulka 19: Překrytí RBF pro sociální model [zdroj: vlastní]

Počet neuronů ve skryté vrstvě	Persistence	$\alpha$	Překrytí RBF vrstvy	Počet cyklů	Přesnost klasifikace na trénovacích datech v %	Přesnost klasifikace na testovacích datech v %
180	30	0,9	0,95	500	97,826	75,194
180	30	0,9	1	500	97,205	75,969
180	30	0,9	2	500	87,267	74,419
180	30	0,9	3	500	83,851	70,543

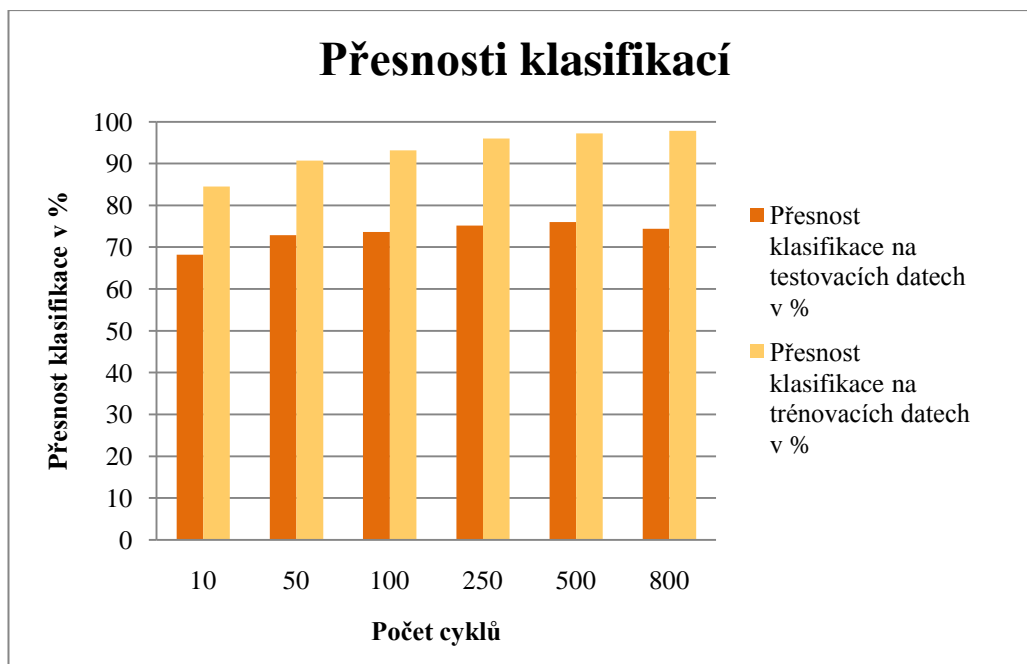


**Graf 13: Přesnosti klasifikací při změně překrytí RBF pro sociální model [zdroj: vlastní]**

Poslední parametr byl počet cyklů. Přehled nastavení počtu cyklů je v následující tabulce 20. Tabulka popisuje, jak se se změnou počtu cyklů měnily i přesnosti klasifikací neuronové sítě. Výsledky jsou znázorněny na grafu 14. Z výsledků vyplývá, že optimální počet cyklů je 500, při nichž přesnosti klasifikací byly na trénovacích datech 97,205% a na testovacích datech 75,969%.

**Tabulka 20: Počet cyklů pro sociální model [zdroj: vlastní]**

Počet neuronů ve skryté vrstvě	Persistence	$\alpha$	Překrytí RBF vrstvy	Počet cyklů	Přesnost klasifikace na trénovacích datech v %	Přesnost klasifikace na testovacích datech v %
180	30	0,9	1	10	84,472	68,217
180	30	0,9	1	50	90,683	72,868
180	30	0,9	1	100	93,168	73,643
180	30	0,9	1	250	95,963	75,194
180	30	0,9	1	500	97,205	75,969
180	30	0,9	1	800	97,826	74,419



Graf 14: Přesnost klasifikace při změně počtů cyklů pro sociální model [zdroj: vlastní]

Prostřednictvím analýzy bylo nakonec nastaveno optimální nastavení neuronové sítě RBF pro sociální model. Optimální nastavení je znázorněno v tabulce 21. Přesnosti klasifikací na trénovacích datech byly 97,205% a na testovacích datech 75,969%.

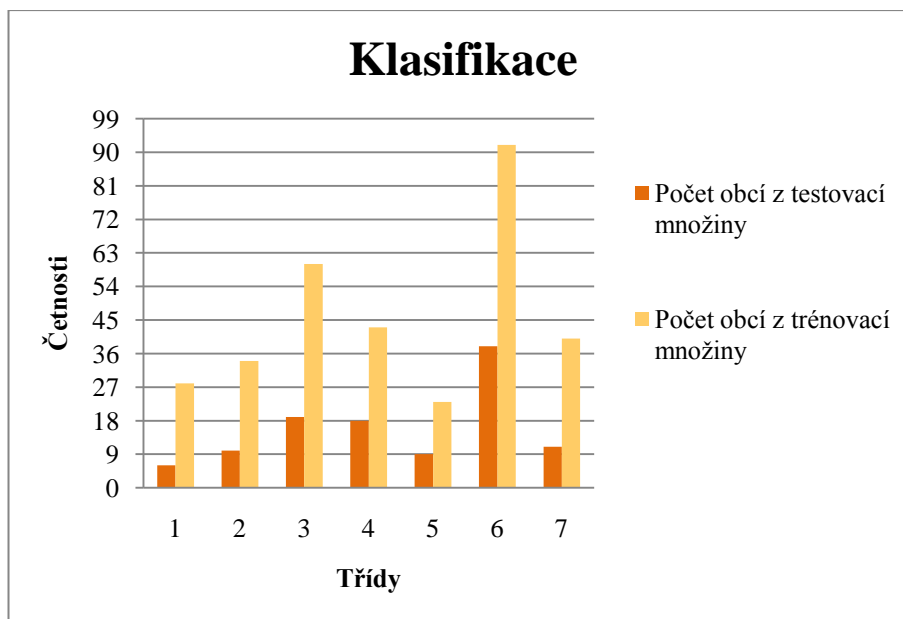
Tabulka 21: Optimální nastavení pro sociální model [zdroj: vlastní]

Počet neuronů ve skryté vrstvě	Persistence	$\alpha$	Překrytí RBF vrstvy	Počet cyklů	Přesnost klasifikace na trénovacích datech v %	Přesnost klasifikace na testovacích datech v %
180	30	0,9	1	500	97,205	75,969

V následující tabulce 22 je přehled výsledné klasifikace obcí. Obce z trénovací a testovací množiny jsou klasifikovány do jednotlivých tříd. Pro lepší přehled slouží graf 15, ze kterého vyplývá, že nejvíc obcí se nachází v 6. třídě. Do této třídy připadlo z trénovací množiny 92 obcí a z testovací 38 obcí. Nejméně obcí z trénovací množiny připadlo do 5. třídy, kam připadlo 23 obcí. Nejméně obcí z testovací množiny připadlo do 1. třídy, kam připadlo 6 obcí.

Tabulka 22: Klasifikace sociálního modelu [zdroj: vlastní]

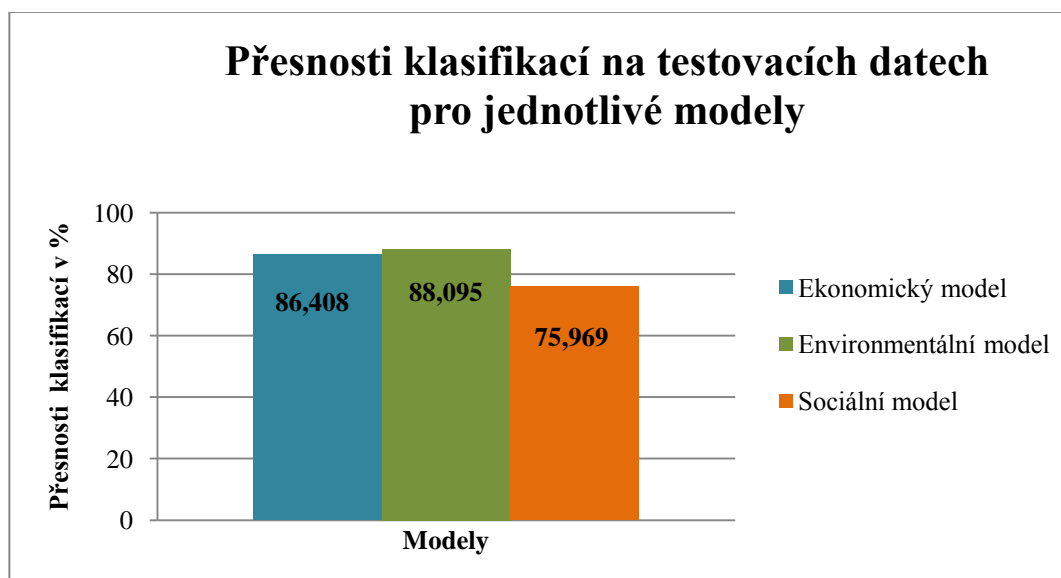
Třídy	1	2	3	4	5	6	7
Počet obcí z testovací množiny	6	10	19	18	9	38	11
Počet obcí z trénovací množiny	28	34	60	43	23	92	40



Graf 15: Klasifikace sociálního modelu [zdroj: vlastní]

#### 4.4 Výsledky klasifikace navržených modelů

Následující graf 16 popisuje, jaké byly přesnosti klasifikací na testovacích datech u jednotlivých modelů. U testovacích dat je poznat, jak je systém naučený, proto jsou tyto přesnosti podstatné. Z grafu vyplývá, že nejlepší přesnost klasifikace měl environmentální model, jehož přesnost klasifikace na testovacích datech činila 88,095%. U ekonomického modelu byla přesnost klasifikace na testovacích datech 86,408%. Nejmenší přesnost klasifikace byla u sociálního modelu, kde přesnost testovacích dat dosáhla 75,969%.



Graf 16: Přesnosti klasifikací modelů [zdroj: vlastní]

## 5 Aproximace

Aproximace byla provedena na všech třech modelech. Vstupní vektory ekonomického a environmentálního modelu představují parametry  $x_1$  až  $x_{11}$ , u sociálního modelu představují parametry  $x_1$  až  $x_{20}$ . Vstupní vektor  $P$  má pak tento tvar [22]:

$$P = (x_1, x_2, \dots, x_n)$$

Protože se daný model týká více než jedné obce, má tvar datové matice [22].

$$P = \begin{matrix} & \begin{matrix} x_1 & \dots & x_k & \dots & x_m \end{matrix} \\ \begin{matrix} o_1 \\ \dots \\ o_i \\ \dots \\ o_n \end{matrix} & \begin{matrix} x_{1,1} & \dots & x_{1,k} & \dots & x_{1,m} \\ \dots & \dots & \dots & \dots & \dots \\ x_{i,1} & \dots & x_{i,k} & \dots & x_{i,m} \\ \dots & \dots & \dots & \dots & \dots \\ x_{n,1} & \dots & x_{n,k} & \dots & x_{n,m} \end{matrix} \end{matrix} \quad (26)$$

Obce v Pardubickém kraji jsou popsány vstupními parametry, které jsou popsány v kapitole 3.1. Míru závislosti mezi proměnnými definujeme pomocí koeficientu korelace [16].

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_x * \sigma_y} \quad (27)$$

- kde:
- $X, Y$  jsou proměnné;
  - $\text{cov}(X, Y)$  je druhý smíšený centrální moment nazývaný kovariance;
  - $\sigma_x, \sigma_y$  jsou výběrové směrodatné odchylky.

Korelační koeficient leží v rozmezí do -1 do 1. Pokud  $\rho(X, Y) > 0$ , jedná se o pozitivně korelované veličiny. Když  $\rho(X, Y) < 0$ , pak jde o negativně korelované veličiny  $|\rho(X, Y)| = 1$  existuje mezi  $X, Y$  vztah, když  $\rho(X, Y) = 0$  jsou  $X$  a  $Y$  navzájem nekorelované (nezávislé). Korelaci lze vypočítat v prostředí MS Excel pomocí funkce CORREL [16].

Vypočítané korelační matice pro jednotlivé modely jsou v příloze 2. I když mezi některými parametry vznikla korelace, bylo expertem doporučeno použít všechny jím navržené parametry. Z toho důvodu byly všechny parametry zahrnuty při nastavování

optimálního nastavení neuronové sítě RBF. Neuronová síť v optimálním nastavení dosahuje uspokojivých přesností klasifikací na trénovacích a na testovacích datech. Cílem aproximace bylo zjistit, jaké parametry ovlivňují přesnost klasifikace při optimálním nastavení neuronové sítě RBF a podat návrh na parametry, které ovlivňují pozitivně přesnosti klasifikací při jejich odebrání.

## 5.1 Ekonomický model

Cílem aproximace bylo zjistit, jaké ekonomické parametry mají vliv na přesnost klasifikace, když neuronová síť RBF je v optimálním nastavení. O optimálním nastavení vypovídá kapitola 4.

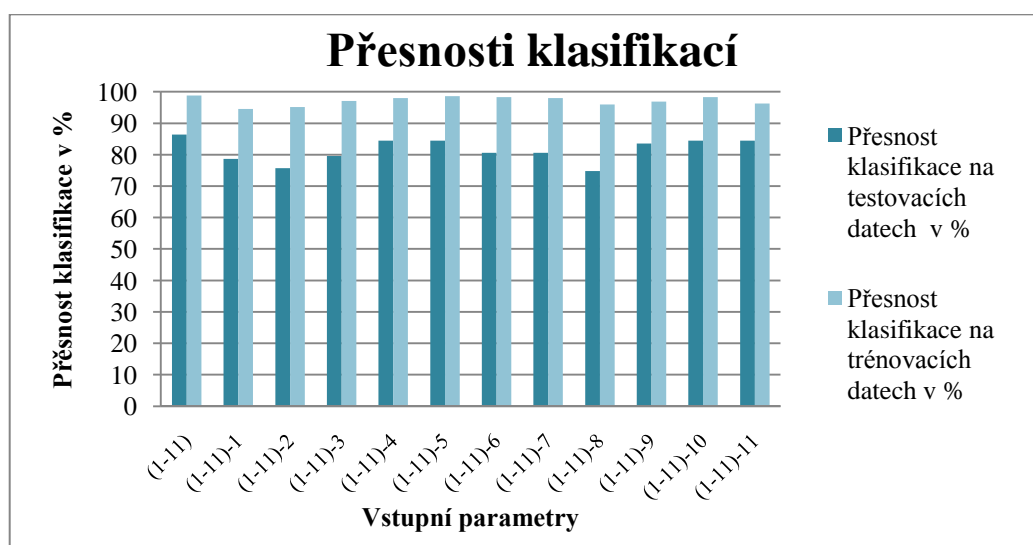
Nejprve bylo zjištěno, jak je přesnost klasifikace ovlivněna při odebrání jednotlivých parametrů (viz. tabulka 23). Tyto přesnosti byly porovnávány s původní přesností klasifikace. Tato analýza byla prováděna jak na trénovacích datech, tak na testovacích datech.

Odchyly přesnosti klasifikace pro trénovací data se pohybují v intervalu  $\langle -4,310; -0,287 \rangle$ . Největší vliv na přesnost trénovacích dat měly parametry EKO1 a EKO2, při jejichž odebrání se přesnost klasifikace trénovacích dat zhoršila.

U testovacích dat se odchylka přesnosti klasifikace pohybovala v intervalu  $\langle -11,650; -1,942 \rangle$ . Ani u nich nedošlo po odebrání jednotlivých parametrů ke zlepšení přesnosti klasifikace. Největší negativní vliv na přesnost klasifikace měly parametry EKO8, EKO2 a EKO1. Pro lepší znázornění slouží graf 20.

Tabulka 23: Přesnosti klasifikací ekonomického modelu parametrů- aproximace [zdroj: vlastní]

Vstupní parametry	Přesnost klasifikace na trénovacích datech v %	Přesnost klasifikace na testovacích datech v %	Odchylka pro trénovací data	Odchylka pro testovací data
<b>(1-11)</b>	<b>98,851</b>	<b>86,408</b>	<b>základní hodnota</b>	
(1-11)-1	94,540	78,641	-4,310	-7,767
(1-11)-2	95,115	75,728	-3,736	-10,680
(1-11)-3	97,126	79,612	-1,724	-6,796
(1-11)-4	97,989	84,466	-0,862	-1,942
(1-11)-5	98,563	84,466	-0,287	-1,942
(1-11)-6	98,276	80,583	-0,575	-5,825
(1-11)-7	97,989	80,583	-0,862	-5,825
(1-11)-8	95,977	74,757	-2,874	-11,650
(1-11)-9	96,839	83,495	-2,011	-2,913
(1-11)-10	98,276	84,466	-0,575	-1,942
(1-11)-11	96,264	84,466	-2,586	-1,942



Graf 17: Přesnosti klasifikací ekonomického modelu parametrů- aproximace [zdroj: vlastní]

Z analýzy, která byla provedena na ekonomickém modelu vyplývá, že žádný z parametrů neovlivnil pozitivně přesnosti klasifikací. Z tohoto důvodu nebyl podán návrh na odebrání parametrů. Všechny ekonomické parametry, které byly navrženy expertem, jsou tedy vhodné pro klasifikaci.

Pro ověření modelu jsou zavedena data validační. Data byla rozdělena takto: 73% trénovací, 20% testovací a 7% validační. Neuronová síť zůstala v optimálním nastavení. Přesnosti klasifikací činily na trénovacích datech 98,193%, na testovacích datech 79,348% a na validačních datech 88,889% (viz. tabulka 24).

Tabulka 24: Verifikace ekonomického modelu [zdroj: vlastní]

Vstupní parametry	Počet neuronů ve skryté vrstvě	Persistence	$\alpha$	Překrytí RBF vrstvy	Počet cyklů	Přesnost klasifikace na trénovacích datech v %	Přesnost klasifikace na testovacích datech v %	Přesnost klasifikace na validačních datech v %
(1-11)	180	30	0,9	1	500	98,193	79,348	88,889

## 5.2 Environmentální model

Stejná analýza aproximace byla provedena u environmentálního modelu. I u něho bylo zkoumáno, jaké parametry mají vliv na přesnost klasifikace. Při těchto analýzách zůstala neuronová síť RBF v optimálním nastavení.

Nejprve bylo zjišťováno, jak jsou přesnosti klasifikací ovlivněny při odebrání jednotlivých parametrů (viz. tabulka 25).

Odchytky přesnosti klasifikace pro trénovací data se pohybují v intervalu  $< -3,704; 0,617 >$ . Pozitivní vliv na přesnost trénovacích dat měly parametry EVO8

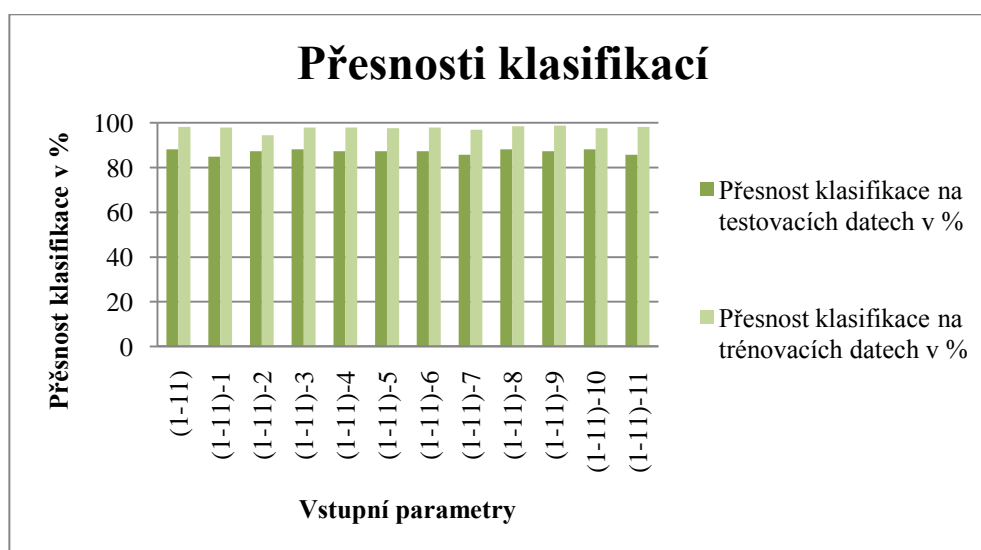


a EVO9, při jejichž odebrání se přesnosti klasifikace trénovacích dat zlepšily. Největší vliv na přesnost měl parametr EVO2, u kterého se přesnost klasifikace trénovacích dat při odebrání zhoršila.

U testovacích dat se odchylka přesnosti klasifikace pohybovala v intervalu  $\langle -3,175;0 \rangle$ . Negativní vliv na přesnost klasifikací měl po odebrání nejvíce parametr EVO1. Při odebrání parametrů EVO3, EVO8 a EVO10 nedošlo u přesnosti klasifikace na testovacích datech k žádným změnám. Z tohoto důvodu bylo usouzeno, že tyto parametry jsou vhodnými adepty na vyřazení. Protože jen u parametru EVO8 došlo ke zlepšení i trénovacích dat, byl navržen na vyřazení on. Pro lepší znázornění slouží graf 18.

Tabulka 25: Přesnosti klasifikací environmentálního modelu- aproximace [zdroj: vlastní]

Vstupy	Přesnost klasifikace na trénovacích datech v %	Přesnost klasifikace na testovacích datech v %	Odchylka pro trénovací data	Odchylka pro testovací data
(1-11)	98,148	88,095	základní hodnota	
(1-11)-1	97,840	84,921	-0,309	-3,175
(1-11)-2	94,444	87,302	-3,704	-0,794
(1-11)-3	97,840	88,095	-0,309	0,000
(1-11)-4	97,840	87,302	-0,309	-0,794
(1-11)-5	97,531	87,302	-0,617	-0,794
(1-11)-6	97,840	87,302	-0,309	-0,794
(1-11)-7	96,914	85,714	-1,235	-2,381
(1-11)-8	98,457	88,095	0,309	0,000
(1-11)-9	98,765	87,302	0,617	-0,794
(1-11)-10	97,531	88,095	-0,617	0,000
(1-11)-11	98,148	85,714	0,000	-2,381



Graf 18: Přesnosti klasifikací environmentálního modelu- aproximace [zdroj: vlastní]

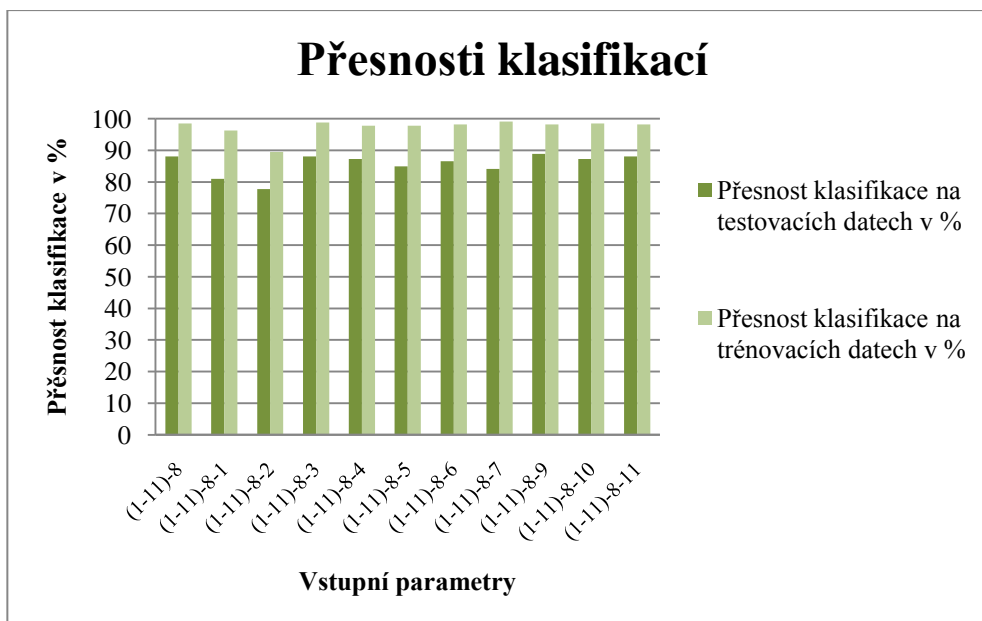
Z předchozí analýzy byl vyřazen parametr EVO8. Následně bylo zkoumáno, jak je přesnost klasifikace ovlivněna při odebrání zbylých parametrů. Tyto přesnosti klasifikací byly porovnávány s přesností klasifikací po odebrání parametru EVO8 (viz. tabulka 26).

Odchytky přesnosti klasifikace pro trénovací data se pohybují v intervalu  $< -8,951; 0,617 >$ . Pozitivní vliv na přesnost trénovacích dat měly parametry EVO7 a EVO3, při jejichž odebrání se přesnosti klasifikací trénovacích dat zlepšily. Nejvíce byla přesnost klasifikace u trénovacích dat opět ovlivněna negativně při odebrání parametru EVO2.

U testovacích dat se odchylka přesnosti klasifikace pohybovala v intervalu  $< -10,317; 0,794 >$ . Nevětší vliv na přesnost klasifikace po odebrání měl parametr EVO2, při jehož odebrání byla přesnost klasifikace testovacích dat negativně ovlivněna. Pozitivně byla přesnost testovacích dat ovlivněna při odebrání parametru EVO9. Pro lepší znázornění slouží graf 19.

**Tabulka 26: Přesnosti klasifikací po odebrání parametru EVO8 [zdroj: vlastní]**

Vstupy	Přesnost klasifikace na trénovacích datech v %	Přesnost klasifikace na testovacích dat v %	Odchytky pro trénovací data	Odchytky pro testovací data
<b>(1-11)-8</b>	<b>98,457</b>	<b>88,095</b>	<b>základní hodnota</b>	
(1-11)-8-1	96,296	80,952	-2,160	-7,143
(1-11)-8-2	89,506	77,778	-8,951	-10,317
(1-11)-8-3	98,765	88,095	0,309	0,000
(1-11)-8-4	97,840	87,302	-0,617	-0,794
(1-11)-8-5	97,840	84,921	-0,617	-3,175
(1-11)-8-6	98,148	86,508	-0,309	-1,587
(1-11)-8-7	99,074	84,127	0,617	-3,968
(1-11)-8-9	98,148	88,889	-0,309	0,794
(1-11)-8-10	98,457	87,302	0,000	-0,794
(1-11)-8-11	98,148	88,095	-0,309	0,000



**Graf 19: Přesnosti klasifikací po odebrání parametru EVO8 [zdroj: vlastní]**

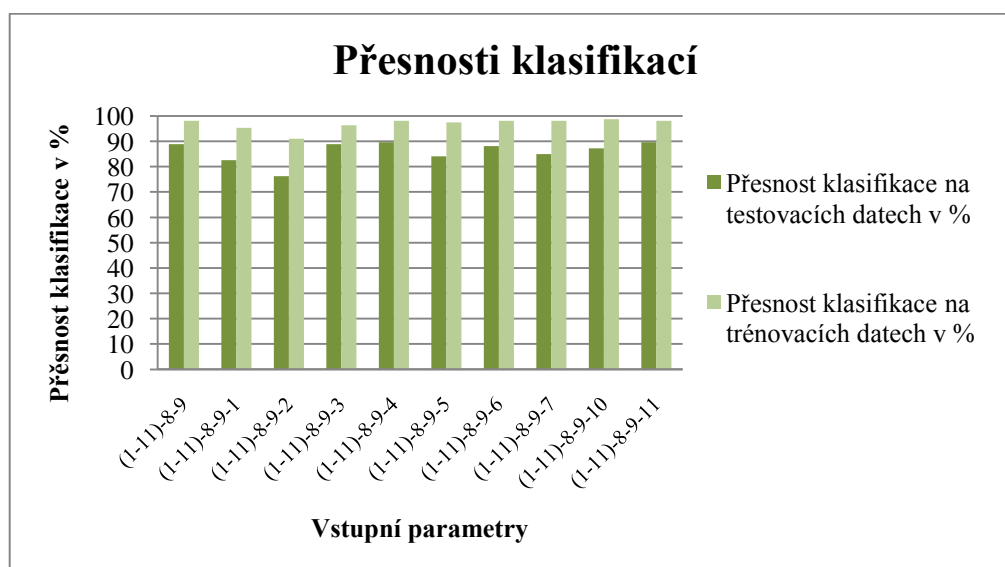
Důležité je, jak se systém naučil, proto byl dalším parametrem na vyřazení označen parametr EVO9. Po odebrání parametrů EVO8 a EVO 9 bylo pokračováno ve zkoumání přesnosti klasifikací při odebrání zbylých parametrů. Tyto přesnosti klasifikací byly porovnávány s přesností klasifikací při odebrání parametrů EVO8 a EVO9 (viz. tabulka 27).

Odchytky přesnosti klasifikace pro trénovací data se pohybují v intervalu  $< -7,099; 0,617 >$ . Pozitivní vliv na přesnost trénovacích dat měl parametr EVO10, při jeho odebrání se přesnosti klasifikace trénovacích dat zlepšily. Nejvíce negativně byla přesnost klasifikace na trénovacích datech opět ovlivněna při odebrání parametru EVO2.

Na testovacích datech se odchylka přesnosti klasifikace pohybovala v intervalu  $< -12,698; 0,794 >$ . Největší vliv na přesnost klasifikace mělo odebrání EVO2. Při jeho odebrání byla přesnost klasifikace na testovacích datech negativně ovlivněna. Pozitivně byla přesnost na testovacích datech ovlivněna při odebrání parametrů EVO4 a EVO11. Pro lepší znázornění slouží graf 20.

Tabulka 27: Přesnosti klasifikací po odebrání parametrů EVO8 a EVO9 [zdroj: vlastní]

Vstupy	Přesnost klasifikace na trénovacích datech v %	Přesnost klasifikace na testovacích datech v %	Odchylka pro trénovací data	Odchylka pro testovací data
<b>(1-11)-8-9</b>	<b>98,148</b>	<b>88,889</b>	<b>základní hodnota</b>	
(1-11)-8-9-1	95,370	82,540	-2,778	-6,349
(1-11)-8-9-2	91,049	76,190	-7,099	-12,698
(1-11)-8-9-3	96,296	88,889	-1,852	0,000
(1-11)-8-9-4	98,148	89,683	0,000	0,794
(1-11)-8-9-5	97,531	84,127	-0,617	-4,762
(1-11)-8-9-6	98,148	88,095	0,000	-0,794
(1-11)-8-9-7	98,148	84,921	0,000	-3,968
(1-11)-8-9-10	98,765	87,302	0,617	-1,587
(1-11)-8-9-11	98,148	89,683	0,000	0,794

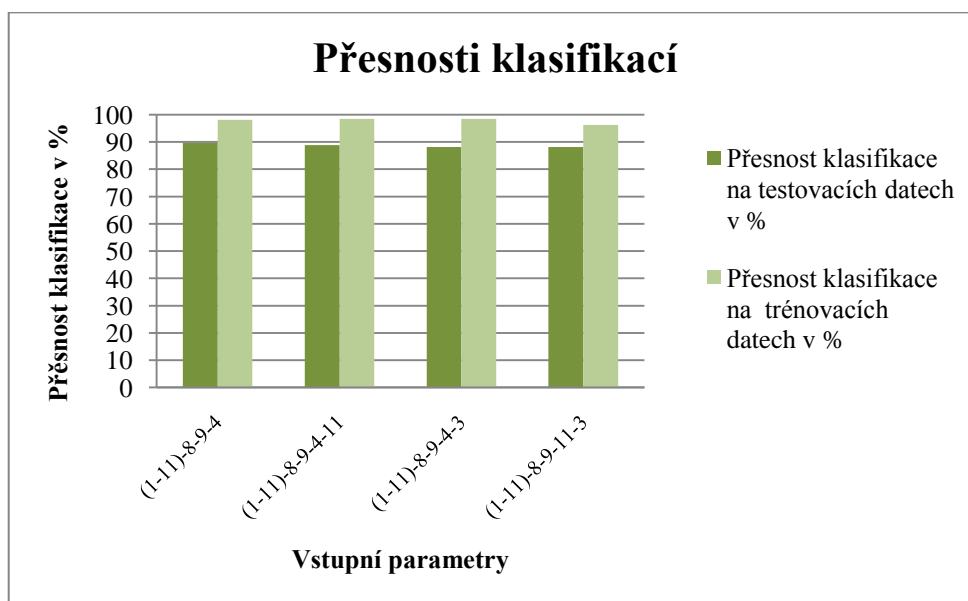


Graf 20: Přesnosti klasifikací po odebrání parametrů EVO8 a EVO9 [zdroj: vlastní]

Při odebrání parametrů EVO4 a EVO11 došlo ke stejnému zlepšení na testovacích datech a oba parametry mají nulový vliv na trénovací data. Proto dále bylo zkoumáno, jak jsou ovlivněny přesnosti klasifikací při odebrání těchto parametrů současně. Při jejich odebrání došlo ke zhoršení přesnosti klasifikace na testovacích datech. Kromě toho byl zkoumán také parametr EVO3, který předchozí analýzu neovlivnil přesnosti klasifikací. Při odebrání EVO3 nedošlo ke zlepšení přesnosti na testovacích datech (viz. tabulka 27).

Tabulka 28: Přesnosti klasifikací po odebrání parametrů EVO8, EVO9, EVO4, EVO11 a EVO3 [zdroj: vlastní]

Vstupy	Přesnost klasifikace na trénovacích datech v %	Přesnost klasifikace na testovacích datech v %	Odchylna pro trénovací data	Odchylna pro testovací data
(1-11)-8-9-4	98,148	89,683	<b>základní hodnota</b>	
(1-11)-8-9-4-11	98,457	88,889	0,309	-0,794
(1-11)-8-9-4-3	98,457	88,095	0,309	-1,587
(1-11)-8-9-11-3	96,296	88,095	-1,852	-1,587



Graf 21: Přesnosti klasifikací po odebrání parametrů EVO8, EVO9, EVO4, EVO11 a EVO3 [zdroj: vlastní]

Z celé analýzy vyplývá, že přesnost klasifikace pro testovací data je nejlepší při odebrání parametrů EVO8, EVO9 a EVO4 či EVO11, které mají stejný vliv na přesnost. Proto byly pro tento model podány dva návrhy. V prvním návrhu jsou vyřazeny parametry EVO8, EVO9 a EVO4. V druhém jsou vyřazeny parametry EVO8, EVO9 a EVO11. Oba návrhy podávají stejné výsledky.

Pro ověření modelu jsou zavedena data validační (viz. tabulka 29). Data byla rozdělena takto: 73% trénovací, 20% testovací a 7% validačních. Neuronová síť zůstala v optimálním nastavení. U prvního návrhu činily přesnosti klasifikací na trénovacích datech 98,093%, na testovacích datech 89,535% a na validačních datech 94,118%. U druhého návrhu činily přesnosti klasifikací na trénovacích datech 98,403%, na testovacích datech 87,535% a na validačních datech 94,118%. Z tohoto důvodu bylo přikloněno k první variantě, kdy přesnost testovacích dat je vyšší.

**Tabulka 29: Verifikace environmentálního modelu [zdroj: vlastní]**

Vstupy	Počet neuronů ve skryté vrstvě	Persistence	$\alpha$	Překrytí RBF vrstvy	Počet cyklů	Přesnost klasifikace na trénovacích datech v %	Přesnost klasifikace na testovacích datech v %	Přesnost klasifikace na validačních datech v %
(1-11)-8-9-4	170	30	0,95	2	500	98,083	89,535	94,118
(1-11)-8-9-11	170	30	0,95	2	500	98,403	87,209	94,118

### 5.3 Sociální model

Také u něho byla provedena stejná analýza aproximace, jaká byla provedena u předcházejících modelů. Bylo zkoumáno, jaké parametry mají vliv na přesnost klasifikace, s cílem nalézt parametry, které přesnost klasifikace po odebrání pozitivně ovlivnily.

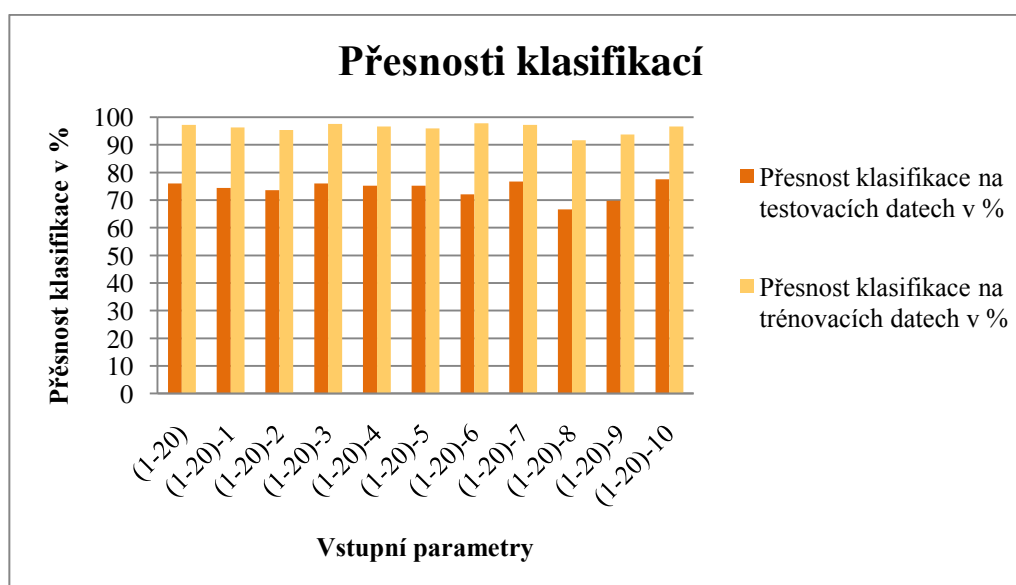
Bylo postupováno stejně jako u předchozích modelů, takže nejprve bylo zjištěno, jak jsou přesnosti klasifikací ovlivněny při odebrání jednotlivých parametrů (viz. tabulka 30).

Odchytky přesnosti klasifikace na trénovacích datech se pohybují v intervalu  $\langle -5,704; 0,932 \rangle$ . Pozitivní vliv na přesnost trénovacích dat měly parametry SOC17 a SOC18, při jejichž odebrání se přesnosti klasifikací trénovacích dat zlepšily. Největší vliv na přesnost měl parametr SOC8, po jehož odebrání došlo ke zhoršení přesnosti klasifikace trénovacích dat.

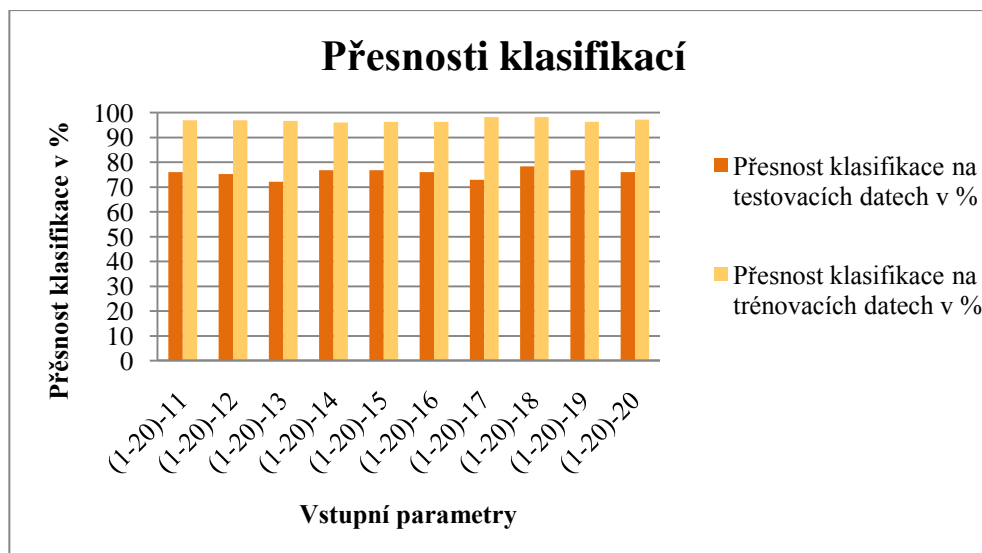
Na testovacích datech se odchylka přesnosti klasifikace pohybovala v intervalu  $\langle -9,302; 2,326 \rangle$ . Negativní vliv na přesnost klasifikací měl po odebrání nejvíce parametr SOC8, tento parametr má po odebrání také největší vliv na přesnost. K nejlepšímu zlepšení přesnosti klasifikací na testovacích datech dochází při odebrání parametru SOC18. Proto byl tento parametr vybrán k vyřazení. Pro lepší znázornění slouží grafy 22 a 23.

Tabulka 30: Aproximace sociálních parametrů [zdroj: vlastní]

Vstupy	Přesnost klasifikace na trénovacích datech v %	Přesnost klasifikace na testovacích datech v %	Odchylna pro trénovací data	Odchylna pro testovací data
(1-20)	97,205	75,969	základní hodnota	
(1-20)-1	96,273	74,419	-0,932	-1,550
(1-20)-2	95,342	73,643	-1,863	-2,326
(1-20)-3	97,516	75,969	0,311	0,000
(1-20)-4	96,584	75,194	-0,621	-0,775
(1-20)-5	95,963	75,194	-1,242	-0,775
(1-20)-6	97,826	72,093	0,621	-3,876
(1-20)-7	97,205	76,744	0,000	0,775
(1-20)-8	91,615	66,667	-5,590	-9,302
(1-20)-9	93,789	69,767	-3,416	-6,202
(1-20)-10	96,584	77,519	-0,621	1,550
(1-20)-11	96,894	75,969	-0,311	0,000
(1-20)-12	96,894	75,194	-0,311	-0,775
(1-20)-13	96,584	72,093	-0,621	-3,876
(1-20)-14	95,963	76,744	-1,242	0,775
(1-20)-15	96,273	76,744	-0,932	0,775
(1-20)-16	96,273	75,969	-0,932	0,000
(1-20)-17	98,137	72,868	0,932	-3,101
(1-20)-18	98,137	78,295	0,932	2,326
(1-20)-19	96,273	76,744	-0,932	0,775
(1-20)-20	97,205	75,969	0,000	0,000



Graf 22: Aproximace sociálních parametrů - první část [zdroj: vlastní]



**Graf 23: Aproximace sociálních parametrů - druhá část [zdroj: vlastní]**

Z předchozí analýzy byl vyřazen parametr SOC18, při jehož odebrání dochází ke zlepšení testovacích dat. Následně bylo zkoumáno, jak je přesnost klasifikace ovlivněna při odebrání zbylých parametrů. Tyto přesnosti klasifikací byly porovnávány s přesností klasifikací po odebrání parametru EVO8 (viz. tabulka 31).

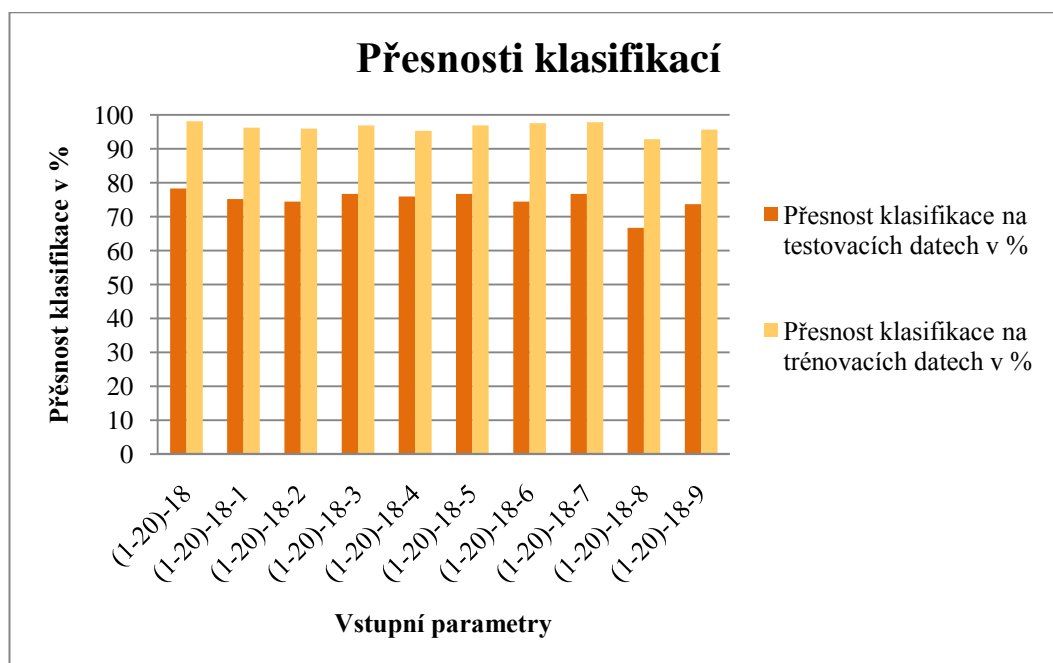
Odchytky přesnosti klasifikace pro trénovací data se pohybují v intervalu  $\langle -5,302; 0,306 \rangle$ . Pozitivní vliv na přesnost klasifikace na trénovacích datech mělo odebrání parametrů SOC19 a SOC20, při jejichž odebrání se přesnosti klasifikací na trénovacích datech zlepšily. Nejvíce byla přesnost klasifikace na trénovacích datech opět ovlivněna negativně při odebrání parametru SOC8.

Na testovacích datech se odchylka přesnosti klasifikace pohybovala v intervalu  $\langle -13,953; 0,775 \rangle$ . Největší vliv na přesnost klasifikace mělo odebrání parametru SOC13, při jehož odebrání byla přesnost klasifikace na testovacích datech negativně ovlivněna. Pozitivně byla přesnost na testovacích datech ovlivněna při odebrání parametru SO12 a SOC19. Z těchto parametrů byl vybrán na vyřazení parametr SOC 19, protože u něho došlo také ke zlepšení trénovacích dat. Pro lepší znázornění slouží grafy 24 a 25.

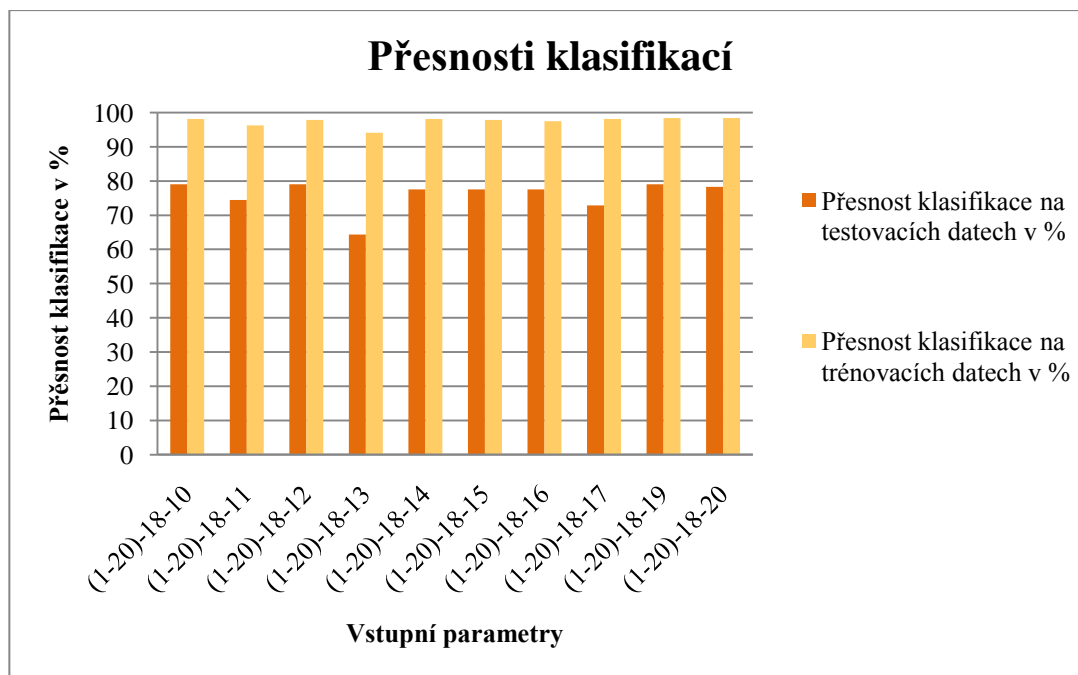


Tabulka 31: Přesnosti klasifikací po odebrání parametru SOC20 [zdroj: vlastní]

Vstupy	Přesnost klasifikace na trénovacích datech v %	Přesnost klasifikace na testovacích datech v %	Odchylka pro trénovací data	Odchylka pro testovací data
<b>(1-20)-18</b>	<b>98,137</b>	<b>78,295</b>	<b>základní hodnota</b>	
(1-20)-18-1	96,273	75,194	-1,863	-3,101
(1-20)-18-2	95,963	74,419	-2,174	-3,876
(1-20)-18-3	96,885	76,744	-1,252	-1,550
(1-20)-18-4	95,327	75,969	-2,810	-2,326
(1-20)-18-5	96,885	76,744	-1,252	-1,550
(1-20)-18-6	97,508	74,419	-0,629	-3,876
(1-20)-18-7	97,819	76,744	-0,317	-1,550
(1-20)-18-8	92,835	66,667	-5,302	-11,628
(1-20)-18-9	95,639	73,643	-2,498	-4,651
(1-20)-18-10	98,131	79,070	-0,006	0,775
(1-20)-18-11	96,262	74,419	-1,875	-3,876
(1-20)-18-12	97,819	79,070	-0,317	0,775
(1-20)-18-13	94,081	64,341	-4,056	-13,953
(1-20)-18-14	98,131	77,519	-0,006	-0,775
(1-20)-18-15	97,819	77,519	-0,317	-0,775
(1-20)-18-16	97,508	77,519	-0,629	-0,775
(1-20)-18-17	98,131	72,868	-0,006	-5,426
(1-20)-18-19	98,442	79,070	0,306	0,775
(1-20)-18-20	98,442	78,295	0,306	0,000



Graf 24: Přesnosti klasifikací po odebrání parametru SOC18 - první část [zdroj: vlastní]



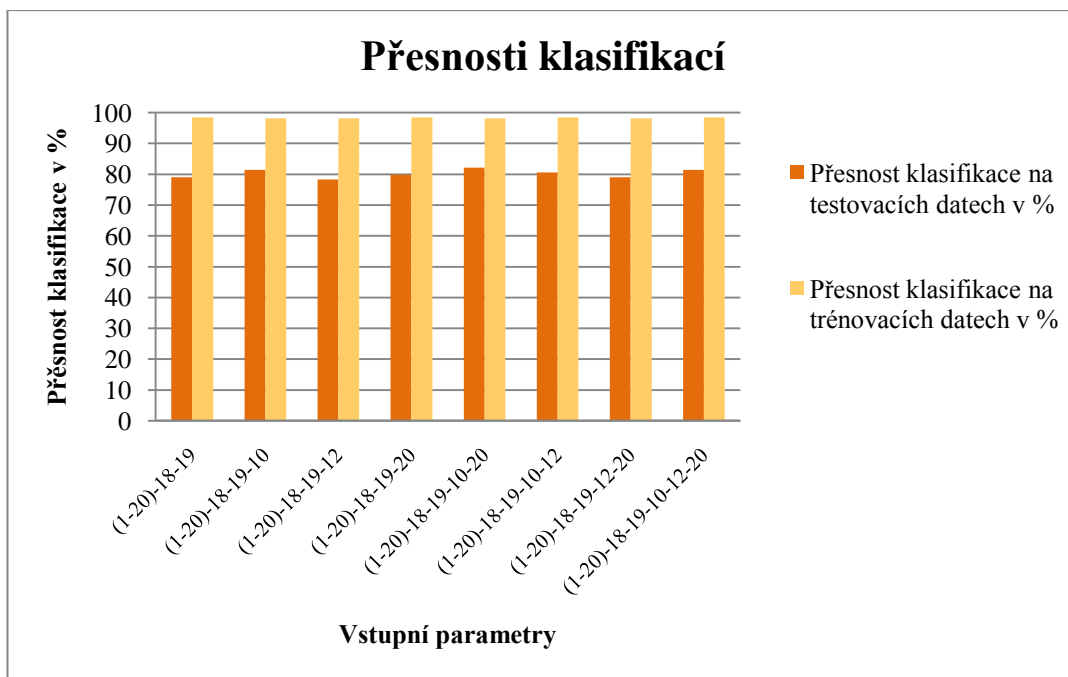
Graf 25: Přesnosti klasifikací po odebrání parametru SOC18- druhá část [zdroj: vlastní]

Sociální model má dvakrát víc parametrů, než modely ekonomické a environmentální, proto bylo rozhodnuto změnit algoritmus, kterým bylo postupováno u předcházejících modelů. Následující analýza byla zaměřena jen na ty parametry, které vykazaly v předchozí analýze po odebrání přesnosti klasifikace u testovacích dat zlepšení (viz. tabulka 32).

Odchyly přesnosti klasifikace pro trénovací data nabývaly hodnot 0 nebo -0,312. U testovacích dat se odchylna přesnosti klasifikace pohybovala v intervalu  $\langle 0; 3,101 \rangle$ . Pro lepší znázornění slouží graf 26.

Tabulka 32: Přesnosti klasifikací po odebrání sociálních parametrů [zdroj: vlastní]

Vstupy	Přesnost klasifikace na trénovacích datech v %	Přesnost klasifikace na testovacích datech v %	Odchylna pro trénovací data	Odchylna pro testovací data
<b>(1-20)-18-19</b>	<b>98,442</b>	<b>79,070</b>	<b>základní hodnota</b>	
(1-20)-18-19-10	98,131	81,395	-0,312	3,101
(1-20)-18-19-12	98,131	78,295	-0,312	0,000
(1-20)-18-19-20	98,442	79,845	0,000	1,550
(1-20)-18-19-10-20	98,131	82,171	-0,312	3,876
(1-20)-18-19-10-12	98,442	80,620	0,000	2,326
(1-20)-18-19-12-20	98,131	79,070	-0,312	0,775
(1-20)-18-19-10-12-20	98,442	81,395	0,000	3,101



Graf 26: Přesnosti klasifikací po odebrání sociálních parametru [zdroj: vlastní]

Z celé analýzy vyplývá, že přesnost klasifikace pro testovací data je nejlepší při odebrání parametrů SOC10 SOC18, SOC19 a SOC20.

Pro ověření modelu jsou zavedena data validační (viz. tabulka 29). Data byla rozdělena takto: 70% trénovací, 23% testovací a 7% validačních. Neuronová síť zůstala v původním nastavení. U prvního návrhu činily přesnosti klasifikací na trénovacích datech 98,408%, na testovacích datech 77,876% a na validačních datech 92%.

Tabulka 33: Verifikace sociálního modelu [zdroj: vlastní]

Vstupy	Počet neuronů ve skryté vrstvě	Persistence	$\alpha$	Překrytí RBF vrstvy	Počet cyklů	Přesnost klasifikace na trénovacích datech v %	Přesnost klasifikace na testovacích datech v %	Přesnost klasifikace na validačních datech v %
(1-17)- 10	180	30	0,9	1	500	98,408	77,876	92,000

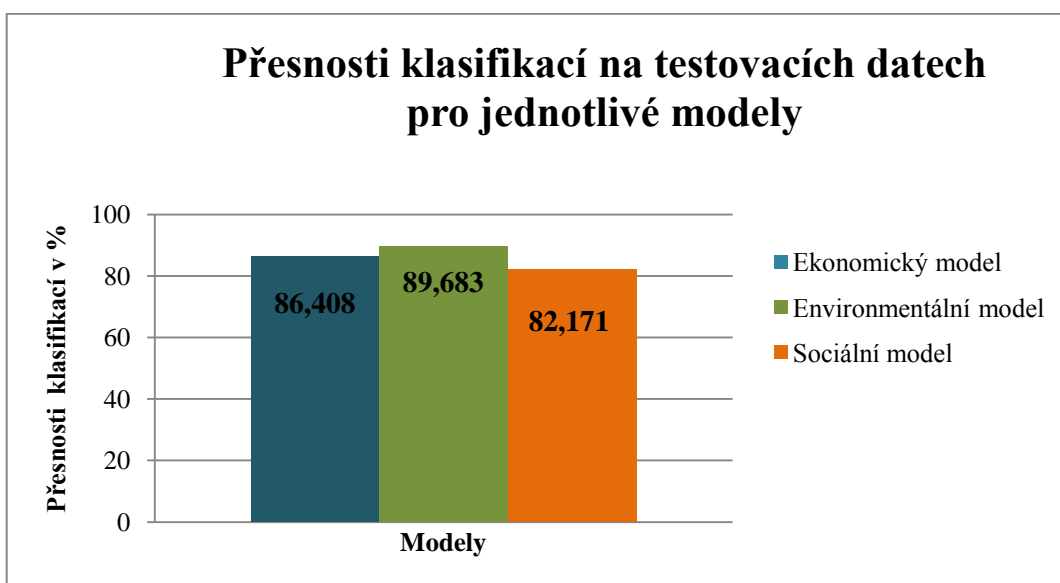
## 5.4 Výsledky aproximace navržených modelů

Parametry pro jednotlivé modely byly navrženy expertem. Úkolem bylo zjistit, zda jsou tyto parametry vhodně navrženy pro klasifikaci. U ekonomického modelu se došlo k závěru, že parametry byly navrženy vhodně pro klasifikaci. Přesnost klasifikace testovacích dat zůstala nezměněna 86,408%.

U environmentálního modelu jsou předloženy dva návrhy na vyřazení parametrů. V prvním návrhu jsou vyřazeny parametry EVO8, EVO9 a EVO4. V druhém jsou vyřazeny parametry EVO8, EVO9 a EVO11. Oba návrhy podávají stejné výsledky, přesnosti klasifikací na testovacích datech se zlepšily z 88,095% na 89,683%.

Na závěr byly zkoumány parametry u sociálního modelu, kde byl podán návrh na vyřazení parametrů SOC10, SOC18, SOC19 a SOC20. U tohoto modelu došlo k velkému zlepšení přesnosti testovacích dat z 75,969% na 82,171%.

Z grafu 27 vyplývá, že nejlepší přesnost klasifikace na testovacích datech měl environmentální model, jehož přesnost klasifikace na testovacích datech činila 89,683%. U ekonomického modelu byla přesnost klasifikace na testovacích datech 86,408%. Nejmenší přesnost klasifikace byla u sociálního modelu, kde přesnost klasifikace na testovacích datech dosáhla 82,171%.



Graf 27: Výsledné přesnosti klasifikací modelů – aproximace [zdroj: vlastní]

## **Závěr**

Cílem diplomové práce byl návrh parametru udržitelného rozvoje pro klasifikaci. Jako metoda byla zvolena neuronová síť typu RBF. Z důvodu, že se udržitelný rozvoj člení na základní tři pilíře, byly vytvořeny tři modely. Parametry pro tyto modely byly navrženy expertem na danou problematiku.

Prvním cílem bylo optimální nastavení neuronové sítě RBF. Při nastavení neuronové sítě byl kladen důraz na to, aby přesnosti klasifikací testovacích dat byly co největší. Podle testovacích dat bylo poznat, jak se systém naučil. Čím vyšší hodnota, tím se systém učil lépe a dokázal tak správně přiřadit obec do třídy.

Po nastavení neuronové sítě do optimálního tvaru bylo zkoumáno, jak jsou přesnosti klasifikací ovlivněny při odebrání jednotlivých parametrů. Cílem byly návrhy na vyřazení parametrů pro jednotlivé modely. Po odebrání těchto parametrů docházelo k zlepšení nebo tyto parametry neměly na přesnosti klasifikací na testovacích datech žádný vliv. Na závěr bylo porovnání jednotlivých modelů, podle jejich přesnosti klasifikací na testovacích datech.

## Seznam obrázků

Obrázek 1: Základní pilíře udržitelného rozvoje [zdroj: vlastní] .....	11
Obrázek 2: Neuron [12] .....	17
Obrázek 3: Model McCulloch-Pittsova neuronu [26].....	17
Obrázek 4: Dopředná neuronová síť [25] .....	18
Obrázek 5: Rekurentní neuronová síť [25] .....	18
Obrázek 6: RBF neuronová síť [22] .....	20
Obrázek 7: Učení s učitelem [5] .....	21
Obrázek 8: Vliv množství trénovacích dat a počet neuronů ve skryté vrstvě na trénovací chybu [4] .	23
Obrázek 9: Vliv počtu RBF neuronů na zevšeobecnující chybu [4].....	24
Obrázek 10: Schematické zobrazení F: $A \rightarrow B$ [13] .....	26
Obrázek 11: RBF síť pro klasifikaci [8] .....	27
Obrázek 12: Model udržitelného rozvoje [zdroj: vlastní] .....	28
Obrázek 13: Rozklad množiny objektů na trénovacích a testovacích množiny [13] .....	37

## Seznam tabulek

Tabulka 1: Ekonomické parametry udržitelného rozvoje .....	29
Tabulka 2: Environmentální parametry (EVN) .....	31
Tabulka 3: Sociální parametry (SOC).....	33
Tabulka 4: Počet neuronů ve skryté vrstvě pro ekonomický model [zdroj: vlastní] .....	40
Tabulka 5: Parametr $\alpha$ pro ekonomický model [zdroj: vlastní] .....	41
Tabulka 6: Překrytí RBF pro ekonomický model [zdroj: vlastní] .....	41
Tabulka 7: Počet cyklů pro ekonomický model [zdroj: vlastní] .....	42
Tabulka 8: Persistence pro ekonomický model [zdroj: vlastní] .....	43
Tabulka 9: Optimální nastavení pro ekonomický model [zdroj: vlastní].....	43
Tabulka 10: Klasifikace ekonomického modelu [zdroj: vlastní] .....	44
Tabulka 11: Počet neuronů ve skryté vrstvě pro environmentální model [zdroj: vlastní] .....	45
Tabulka 12: Parametr $\alpha$ pro environmentální model [zdroj: vlastní] .....	45
Tabulka 13: Překrytí RBF pro environmentální model [zdroj: vlastní] .....	46
Tabulka 14: Počet cyklů pro environmentální model [zdroj: vlastní].....	47
Tabulka 15: Optimální nastavení pro environmentální model [zdroj: vlastní] .....	47
Tabulka 16: Klasifikace environmentálního modelu [zdroj: vlastní].....	48
Tabulka 17: Počet neuronů ve skryté vrstvě pro sociální model [zdroj: vlastní] .....	49
Tabulka 18: Parametr $\alpha$ pro sociální model [zdroj: vlastní].....	50
Tabulka 19: Překrytí RBF pro sociální model [zdroj: vlastní].....	50
Tabulka 20: Počet cyklů pro sociální model [zdroj: vlastní] .....	51
Tabulka 21: Optimální nastavení pro sociální model [zdroj: vlastní].....	52
Tabulka 22: Klasifikace sociálního modelu [zdroj: vlastní] .....	52
Tabulka 23: Přesnosti klasifikací ekonomického modelu parametrů- aproximace [zdroj: vlastní] ....	55
Tabulka 24: Verifikace ekonomického modelu [zdroj: vlastní].....	56
Tabulka 25: Přesnosti klasifikací environmentálního modelu- aproximace [zdroj: vlastní].....	57
Tabulka 26: Přesnosti klasifikací po odebrání parametru EVO8 [zdroj: vlastní] .....	58
Tabulka 27: Přesnosti klasifikací po odebrání parametrů EVO8 a EVO9 [zdroj: vlastní].....	60
Tabulka 28: Přesnosti klasifikací po odebrání parametrů EVO8, EVO9, EVO4, EVO11 a EVO3 [zdroj: vlastní].....	61
Tabulka 29: Verifikace environmentálního modelu [zdroj: vlastní] .....	62
Tabulka 30: Aproximace sociálních parametrů [zdroj: vlastní] .....	63
Tabulka 31: Přesnosti klasifikací po odebrání parametru SOC20 [zdroj: vlastní].....	65
Tabulka 32: Přesnosti klasifikací po odebrání sociálních parametrů [zdroj: vlastní] .....	66
Tabulka 33: Verifikace sociálního modelu [zdroj: vlastní].....	67
Tabulka 34: Korelační matice ekonomických parametrů [zdroj: vlastní].....	79
Tabulka 35: Korelační matice environmentálních parametrů [zdroj: vlastní] .....	79
Tabulka 36: Korelační matice sociálních parametrů [zdroj: vlastní] .....	80

## Seznam grafů

Graf 1: Přesnosti klasifikací při změně počtu neuronů ve skryté vrstvě pro ekonomický model [zdroj: vlastní] .....	40
Graf 2: Přesnosti klasifikací při změně parametru $\alpha$ pro ekonomický model [zdroj: vlastní] .....	41
Graf 3: Přesnosti klasifikací při změně překrytí RBF pro ekonomický model [zdroj: vlastní].....	42
Graf 4: Přesnosti klasifikací při změně počtu cyklů pro ekonomický model [zdroj: vlastní] .....	43
Graf 5: Klasifikace ekonomického modelu [zdroj: vlastní] .....	44
Graf 6: Přesnosti klasifikací při změně počtu neuronů ve skryté vrstvě pro environmentální model [zdroj: vlastní].....	45
Graf 7: Přesnosti klasifikací při změně parametru $\alpha$ pro environmentální model [zdroj: vlastní] .....	46
Graf 8: Přesnosti klasifikací při změně překrytí RBF pro environmentální model [zdroj: vlastní] ....	46
Graf 9: Přesnosti klasifikací při změně počtu cyklů pro environmentální model [zdroj: vlastní].....	47
Graf 10: Klasifikace environmentálního modelu [zdroj: vlastní] .....	48
Graf 11: Přesnosti klasifikací při změně počtu neuronů ve skryté vrstvě pro sociální model [zdroj: vlastní] .....	49
Graf 12: Přesnosti klasifikací při změně parametru $\alpha$ pro sociální model [zdroj: vlastní] .....	50
Graf 13: Přesnosti klasifikací při změně překrytí RBF pro sociální model [zdroj: vlastní] .....	51
Graf 14: Přesnost klasifikace při změně počtu cyklů pro sociální model [zdroj: vlastní] .....	52
Graf 15: Klasifikace sociálního modelu [zdroj: vlastní] .....	53
Graf 16: Přesnosti klasifikací modelů [zdroj: vlastní] .....	53
Graf 17: Přesnosti klasifikací ekonomického modelu parametrů- aproximace [zdroj: vlastní] .....	56
Graf 18: Přesnosti klasifikací environmentálního modelu- aproximace [zdroj: vlastní] .....	57
Graf 19: Přesnosti klasifikací po odebrání parametru EVO8 [zdroj: vlastní] .....	59
Graf 20: Přesnosti klasifikací po odebrání parametrů EVO8 a EVO9 [zdroj: vlastní] .....	60
Graf 21: Přesnosti klasifikací po odebrání parametrů EVO8, EVO9, EVO4, EVO11 a EVO3 [zdroj: vlastní] .....	61
Graf 22: Aproximace sociálních parametrů - první část [zdroj: vlastní].....	63
Graf 23: Aproximace sociálních parametrů - druhá část [zdroj: vlastní] .....	64
Graf 24: Přesnosti klasifikací po odebrání parametru SOC18 - první část [zdroj: vlastní].....	65
Graf 25: Přesnosti klasifikací po odebrání parametru SOC18- druhá část [zdroj: vlastní] .....	66
Graf 26: Přesnosti klasifikací po odebrání sociálních parametrů [zdroj: vlastní] .....	67
Graf 27: Výsledné přesnosti klasifikací modelů - aproximace [zdroj: vlastní].....	68



## Seznam příloh

Příloha 1 .....	77
Příloha 2 .....	79

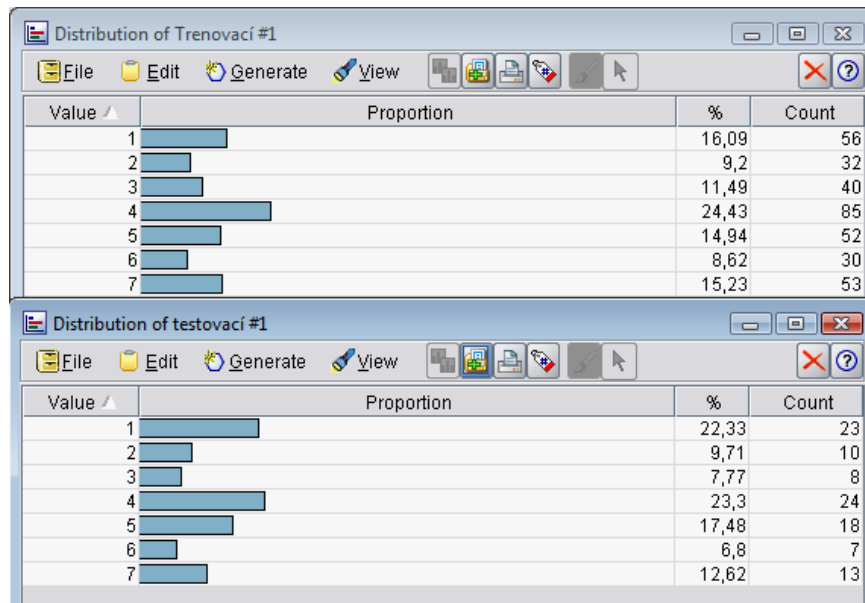
## Použitá literatura

- [1] BORS, A. *Introduction of the Radial basis function (RBF) networks* [online]. 2000 [cit. 2010-02-11]. Dostupný z WWW: <<http://users.cecs.anu.edu.au/~daa/courses/GSAC6017/rbf.pdf>>.
- [2] ČSÚ: *Český statistický úřad*, [online]. 2009 [cit. 2009-10-17]. Dostupný z WWW: <<http://www.czso.cz>>.
- [3] FIALOVÁ, H. *Vybrané oblasti udržitelného rozvoje v Pardubickém kraji*. Pardubice: Český statistický úřad, oddělení regionálních analýz a informačních služeb Pardubice, 2007. ISBN 978-80-250-1621-3.
- [4] GIROSI, F., NIYOGI, P. *On the Relationship Between Generalization Error, Hypothesis Complexity, and Sample Complexity for Radial Basis Functions* [online]. 1994 [cit. 2010-02-12]. Dostupný z WWW: <<ftp://publications.ai.mit.edu/ai-publications/pdf/AIM-1467.pdf>>.
- [5] HAYKIN, S. *Neural Networks: A Comprehensive Foundation*. 2nd edition, New Jersey, Prentice-Hall, Inc., 1999, 842s. ISBN 81-7808-300-0.
- [6] JAHN, M. *Strategie udržitelného rozvoje České republiky*. Praha : Úřad vlády České republiky, 2005. ISBN 80-86734-42-0.
- [7] KELBA, J., ŠILHÁN, D. *Shluková analýza* [online]. 2009 [cit. 2010-02-12]. Dostupný z WWW: <<http://staff.utia.cas.cz/nagy/skola/Projekty/Classification/ShlukovaAnalyza.pdf>>.
- [8] KLÍMA, D. *Neuronové sítě typu RBF* [online]. 2000 [cit. 2010-02-12]. Dostupný z WWW: <<http://www.volny.cz/klimad/skola/spa/skrbek%20web/rbf/rbf.html#uceni>>.
- [9] KOCHANÍČKOVÁ, M. *Modelování bonity obcí pomocí RBF neuronových sítí*. Pardubice, 2008. 67 s. Univerzita Pardubice, Ústav systémového inženýrství a informatiky. Vedoucí diplomové práce Olej V.
- [10] Krajský úřad Pardubického kraje. *Pardubický kraj* [online]. 2008 [cit. 2009-10-10]. Dostupný z WWW: <<http://www.pardubickykraj.cz>>.
- [11] KREBS, V. *Formulace doporučení MPSV v oblasti sociálního pilíře udržitelného rozvoje vycházejících z mezinárodní komparace*. Praha: Projekt MPSV HR139/05 – Závěrečná zpráva, 2005.

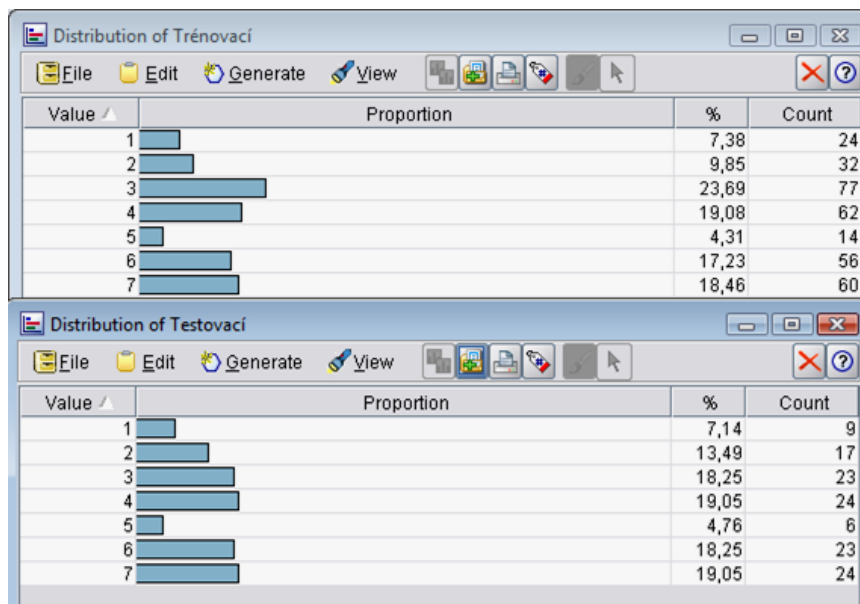
- [12] KUBA, M. *Neuronové sítě* [online]. Brno: Masarykova univerzita, 1995 [cit. 2010-04-01]. Dostupné z WWW: <<http://tf.czu.cz/~votruba/WAN/Neuronove%20site.pdf>>.
- [13] KVASNIČKA, V., et al. *Úvod do teórie neurónových sietí*. Bratislava: Iris, 1997. 285 s. ISBN 80-88778-30-1.
- [14] KVASNIČKA, V., SKLENÁK, Š., POSPÍCHAL, J. *Neuronové sítě pro chemii teorie, software, hardware*. Chemické listy [online]. 1993, 87, [cit. 2010-27-03]. Dostupný z WWW: <[www.jh-inst.cas.cz/~sklenak/stepan/publ/6.pdf](http://www.jh-inst.cas.cz/~sklenak/stepan/publ/6.pdf)>.
- [15] MAŘÍK, V., et al. *Umělá inteligence 4*. Praha: Akademie věd České republiky, 2003. 475 s. ISBN 80-200-1044-0.
- [16] MELOUN, M., MILITKÝ, J. *Kompendium statistického zpracování dat: Metody a řešené úlohy včetně CD*. Praha: Akademie věd České republiky, 2002. 764 s. ISBN 80-200-1008-4.
- [17] MEZŘICKÝ, V. *Environmentální politika a udržitelný rozvoj*. Praha: Portál, 2005. ISBN 80-7367-003-8.
- [18] MOLDAN, B., HÁK, T., KOLÁŘOVÁ, H. *K udržitelnému rozvoji České republiky: vytváření podmínek*. Praha: Centrum Univerzity Karlovy pro životního prostředí, 2002. Sv. IV: Vzdělání, informace, indikátory. ISBN 80-238-8378-X.
- [19] MOLDAN, B. *Příroda a civilizace, životní prostředí a rozvoj lidské civilizace*. Praha: Státní pedagogické nakladatelství, 1997. ISBN 80-04-26434-4.
- [20] NÁTR, L. *Rozvoj trvale neudržitelný*. Praha: Karolinum, 2006. ISBN 80-246-0987-8.
- [21] OLEJ, V. *Modelovanie ekonomického procesov na báze vypočtovej inteligencie*. [Vědecká monografie], Miloš Vognar- M&V, Hradec Králové, Česká republika, 2003. 160s. ISBN 80-903024-9-1.
- [22] OLEJ, V. HÁJEK, P. *Municipal Creditworthiness Modelling by Radial Basis Function Neural Networks and Sensitive Analysis their Inputs Parameters*. 19th International Conference on Artificial Neural Network. ICANN 2009, 14-17 September, Limassol, Cyprus, Alippi, C., Polycarpou, M., Panayiotou, Ch., Ellinas, G., Eds., Springer Berlin Heidelberg New York, 2009, pp.505-514. ISSN 0302-9743, ISBN 3-642-04276-7.

- [23] O NÁS? S NÁMI! aneb strategie udržitelného rozvoje očima aktérů environmentální, ekonomické a sociální sféry [online]. [2007] [cit. 2009-11-20]. Dostupný z WWW: <<http://udrzitelnyrozvoj.ecn.cz/>>.
- [24] ŘEZANKOVÁ, H., HÚSEK, D., SNÁŠEL, V. *Shluková analýza dat*. Praha: Professional Publishing, 2007. 196 s. ISBN 978-80-86946-26-9.
- [25] ŠÍMA, J., NERUDA, R. *Teoretické otázky neuronových sítí* [online]. 1996 [cit. 2010-02-12]. Dostupný z WWW: <<http://www2.cs.cas.cz/~sima/kniha.pdf>>.
- [26] TUČKOVÁ, J. *Vybrané aplikace umělých neuronových sítí při zpracování signálů*. Praha: České vysoké učení techniky v Praze, 2009. 224 s. ISBN 978-80-01-04229-8.
- [27] UNIVERZITA KARLOVA. *Centrum pro otázky životního prostředí - Co jsou to indikátory* [online]. 2009 [cit. 2009-09-25]. Dostupný z WWW: <<http://cozp.cuni.cz/COZP-14.html>>.
- [28] VLČEK, J. *Modelování udržitelného rozvoje obcí pomocí fuzzy logiky*. Pardubice, 2009. 88 s. Univerzita Pardubice, Ústav systémového inženýrství a informatiky. Vedoucí diplomové práce Hájek P.

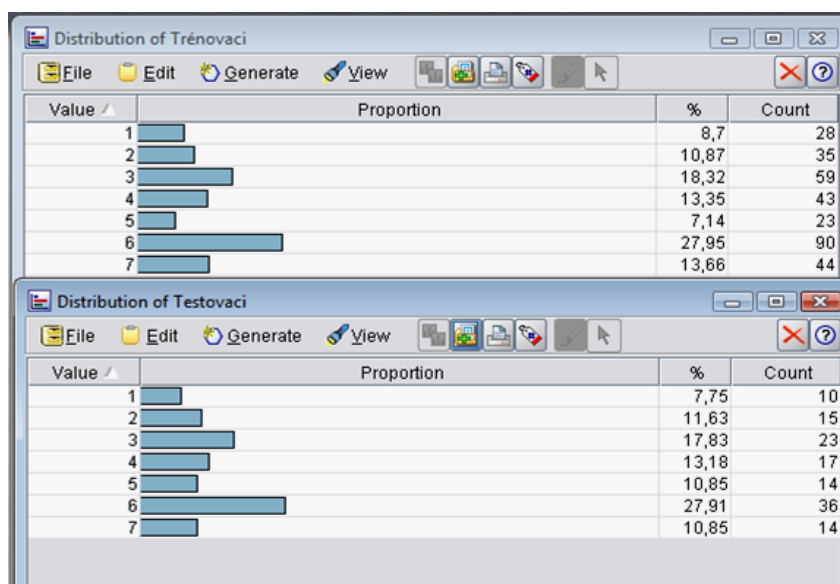
# Příloha 1



Obrázek 14: Rozdělení dat na trénovací a testovací množinu u ekonomického modelu [zdroj: vlastní]



Obrázek 15: Rozdělení dat na trénovací a testovací množinu u environmentálního modelu [zdroj: vlastní]



Obrázek 16: Rozdělení dat na trénovací a testovací množinu u sociálního modelu [zdroj: vlastní]

## Příloha 2

Tabulka 34: Korelační matice ekonomických parametrů [zdroj: vlastní]

Korelace	EKO1	EKO2	EKO3	EKO4	EKO5	EKO6	EKO7	EKO8	EKO9	EKO10	EKO11
EKO1	<b>1</b>	0,739	-0,259	0,029	-0,035	0,126	0,146	-0,005	-0,048	-0,255	-0,205
EKO2	0,739	<b>1</b>	-0,205	0,059	-0,035	-0,005	0,026	0,009	-0,085	-0,072	-0,261
EKO3	-0,259	-0,205	<b>1</b>	0,015	0,06	0,085	-0,071	-0,038	0,003	-0,035	0,298
EKO4	0,029	0,059	0,015	<b>1</b>	0,146	0,02	-0,039	0,109	0,06	0,033	-0,058
EKO5	-0,035	-0,035	0,06	0,146	<b>1</b>	-0,002	0,055	0,064	0,383	0,082	0,038
EKO6	0,126	-0,005	0,085	0,02	-0,002	<b>1</b>	0,094	0,42	-0,125	-0,557	-0,03
EKO7	0,146	0,026	-0,071	-0,039	0,055	0,094	<b>1</b>	-0,005	0,337	-0,18	-0,04
EKO8	-0,005	0,009	-0,038	0,109	0,064	0,42	-0,005	<b>1</b>	0,09	0,035	-0,03
EKO9	-0,048	-0,085	0,003	0,06	0,383	-0,125	0,337	0,09	<b>1</b>	0,222	0,136
EKO10	-0,255	-0,072	-0,035	0,033	0,082	-0,557	-0,18	0,035	0,222	<b>1</b>	0,052
EKO11	-0,205	-0,261	0,298	-0,058	0,038	-0,03	-0,04	-0,03	0,136	0,052	<b>1</b>

Tabulka 35: Korelační matice environmentálních parametrů [zdroj: vlastní]

	EVN1	EVN2	EVN3	EVN4	EVN5	EVN6	EVN7	EVN8	EVN9	EVN10	EVN11
EVN1	<b>1</b>	-0,475	-0,051	0,347	0,17	0,023	-0,335	-0,034	-0,14	-0,125	-0,352
EVN2	-0,475	<b>1</b>	0,077	-0,131	-0,058	-0,089	0,034	0,033	-0,025	-0,049	0,095
EVN3	-0,051	0,077	<b>1</b>	0,036	0,055	-0,724	-0,102	0,061	-0,629	-0,334	0,106
EVN4	0,347	-0,131	0,036	<b>1</b>	0,213	-0,202	-0,373	-0,035	-0,251	-0,297	-0,385
EVN5	0,17	-0,058	0,055	0,213	<b>1</b>	-0,11	-0,212	-0,077	-0,122	-0,176	-0,191
EVN6	0,023	-0,089	-0,724	-0,202	-0,11	<b>1</b>	0,157	-0,041	0,532	0,711	-0,021
EVN7	-0,335	0,034	-0,102	-0,373	-0,212	0,157	<b>1</b>	0,249	0,5	0,434	0,721
EVN8	-0,034	0,033	0,061	-0,035	-0,077	-0,041	0,249	<b>1</b>	-0,009	0,037	0,227
EVN9	-0,14	-0,025	-0,629	-0,251	-0,122	0,532	0,5	-0,009	<b>1</b>	0,576	0,556
EVN10	-0,125	-0,049	-0,334	-0,297	-0,176	0,711	0,434	0,037	0,576	<b>1</b>	0,37
EVN11	-0,352	0,095	0,106	-0,385	-0,191	-0,021	0,721	0,227	0,556	0,37	<b>1</b>

**Tabulka 36: Korelační matice sociálních parametrů [zdroj: vlastní]**

	SOC1	SOC2	SOC3	SOC4	SOC5	SOC6	SOC7	SOC8	SOC9	SOC10	SOC11	SOC12	SOC13	SOC14	SOC15	SOC16	SOC17	SOC18	SOC19	SOC20
SOC1	1	0,499	-0,183	-0,345	0,070	-0,019	-0,099	0,073	-0,232	-0,359	-0,168	-0,165	-0,277	-0,205	-0,113	-0,051	0,018	-0,113	-0,006	0,010
SOC2	0,499	1	-0,117	-0,484	-0,100	0,024	-0,077	-0,118	-0,292	-0,299	0,058	-0,259	-0,345	-0,165	-0,140	-0,084	-0,017	0,012	-0,024	-0,043
SOC3	-0,183	-0,117	1	0,252	0,052	-0,065	0,082	0,106	0,191	0,092	-0,121	0,242	0,124	0,096	0,111	-0,039	-0,051	-0,086	-0,006	0,045
SOC4	-0,345	-0,484	0,252	1	0,008	-0,126	0,015	0,110	0,309	0,178	-0,046	0,257	0,158	0,137	0,072	0,017	-0,112	-0,031	-0,055	0,048
SOC5	0,070	-0,100	0,052	0,008	1	0,497	-0,082	-0,051	-0,026	0,078	0,053	0,090	-0,012	-0,007	-0,001	-0,306	0,635	0,176	-0,097	-0,016
SOC6	-0,019	0,024	-0,065	-0,126	0,497	1	-0,124	-0,351	-0,313	-0,138	0,274	-0,210	-0,224	-0,128	-0,103	-0,175	0,521	0,269	-0,060	-0,040
SOC7	-0,099	-0,077	0,082	0,015	-0,082	-0,124	1	0,088	-0,061	0,034	-0,122	0,043	0,093	0,051	0,013	0,086	-0,057	-0,071	0,027	0,011
SOC8	0,073	-0,118	0,106	0,110	-0,051	-0,351	0,088	1	0,326	0,187	-0,308	0,269	0,098	0,138	0,133	-0,101	-0,091	-0,271	0,161	0,030
SOC9	-0,232	-0,292	0,191	0,309	-0,026	-0,313	-0,061	0,326	1	0,337	-0,228	0,363	0,259	0,203	0,161	-0,003	-0,206	-0,202	0,017	0,055
SOC10	-0,359	-0,299	0,092	0,178	0,078	-0,138	0,034	0,187	0,337	1	-0,085	0,338	0,256	0,707	0,576	-0,044	-0,051	-0,063	0,016	0,022
SOC11	-0,168	0,058	-0,121	-0,046	0,053	0,274	-0,122	-0,308	-0,228	-0,085	1	-0,345	-0,058	-0,082	-0,115	0,027	0,134	0,724	-0,110	-0,050
SOC12	-0,165	-0,259	0,242	0,257	0,090	-0,210	0,043	0,269	0,363	0,338	-0,345	1	0,191	0,247	0,193	-0,062	-0,069	-0,274	0,068	0,146
SOC13	-0,277	-0,345	0,124	0,158	-0,012	-0,224	0,093	0,098	0,259	0,256	-0,058	0,191	1	0,150	0,125	-0,025	-0,136	-0,046	0,059	-0,027
SOC14	-0,205	-0,165	0,096	0,137	-0,007	-0,128	0,051	0,138	0,203	0,707	-0,082	0,247	0,150	1	0,534	0,060	-0,127	-0,049	0,074	-0,006
SOC15	-0,113	-0,140	0,111	0,072	-0,001	-0,103	0,013	0,133	0,161	0,576	-0,115	0,193	0,125	0,534	1	0,011	-0,049	-0,085	0,274	0,022
SOC16	-0,051	-0,084	-0,039	0,017	-0,306	-0,175	0,086	-0,101	-0,003	-0,044	0,027	-0,062	-0,025	0,060	0,011	1	-0,348	0,025	0,098	0,026
SOC17	0,018	-0,017	-0,051	-0,112	0,635	0,521	-0,057	-0,091	-0,206	-0,051	0,134	-0,069	-0,136	-0,127	-0,049	-0,348	1	0,087	-0,096	-0,036
SOC18	-0,113	0,012	-0,086	-0,031	0,176	0,269	-0,071	-0,271	-0,202	-0,063	0,724	-0,274	-0,046	-0,049	-0,085	0,025	0,087	1	-0,026	-0,043
SOC19	-0,006	-0,024	-0,006	-0,055	-0,097	-0,060	0,027	0,161	0,017	0,016	-0,110	0,068	0,059	0,074	0,274	0,098	-0,096	-0,026	1	-0,015
SOC20	0,010	-0,043	0,045	0,048	-0,016	-0,040	0,011	0,030	0,055	0,022	-0,050	0,146	-0,027	-0,006	0,022	0,026	-0,036	-0,043	-0,015	1