

# **Klasifikace audio signálů pomocí neuronové sítě**

Classification of audio signals using  
neural network

**Bc. Veronika Dopitová**

---

Diplomová práce  
2011



Univerzita Tomáše Bati ve Zlíně  
Fakulta aplikované informatiky

---

Univerzita Tomáše Bati ve Zlíně

Fakulta aplikované informatiky

akademický rok: 2010/2011

## ZADÁNÍ DIPLOMOVÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Bc. Veronika DOPITOVÁ**  
Osobní číslo: **A09357**  
Studijní program: **N 3902 Inženýrská informatika**  
Studijní obor: **Bezpečnostní technologie, systémy a management**

Téma práce: **Klasifikace audio signálů pomocí neuronové sítě**

Zásady pro vypracování:

1. Prostudujte doporučenou literaturu a zpracujte literární rešerši na dané téma.
2. Zaměřte se na problematiku neuronových sítí, vhodně zvolte architekturu neuronové sítě pro klasifikaci.
3. Seznamte se se stávající databází zvuků, která bude představovat trénovací množinu.
4. Z uvedené databáze extrahujte charakteristiky (spektrum apod.) které povedou k úspěšné klasifikaci předkládaných zvuků.
5. Využijte vhodné programové prostředí (matlab, mathematica) a vytvořte algoritmus realizující neuronovou síť s možností klasifikace neznámých zdrojů zvuků.
6. Proveďte optimalizaci topologie neuronové sítě.
7. Zhodnoťte dosažené výsledky a navrhněte možnosti pokračování.

Rozsah diplomové práce:

Rozsah příloh:

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

Seznam odborné literatury:

1. Sibbald, A.: An introduction to sound and hearing [online]. 1999 [cit. 2007-02-10]. Dostupný z WWW:<http://sensaaura.com>
2. Kolmer, F., Kyncl, J.: Prostorová akustika. Praha : SNTL – Nakladatelství technické literatury, 1980. 244 s.
3. Frigo, M., Johnson, S. G. FFTW3, [online]. 2006 [cit. 2007-02-15]. Dostupný z WWW: <http://fftw.org>
4. Zelinka Ivan. Umělá inteligence I. Volume 1. Zlín : Vutium, Brno, 1998. 126 p. ISBN 80-214-1163-5.
5. Kvasnička V., Beňušková., Pospíchal J., Farkaš I., Tiňo P. a Král A. (1997). Úvod do teórie neuronových sietí, Iris: Bratislava.
6. Ben Kröse and Patrick van der Smagt (1996), An introduction to Neural Networks.
7. Sinčák , Peter, ANDREJKOVÁ, Gabriela. Neuronové siete Inžiniersky prístup [online]. Dostupný z WWW: <http://hq.alert.sk/mandos/fmfi-uk/Informatika/Neuronove%20Siete/NS1/node1.html>.
8. Beňušková L. Umele neuronové siete. In: Návrat P., Bieliková M., Benuskova L, Kapustik I, Unger M. Umela inteligencia

Vedoucí diplomové práce:

**Ing. Milan Navrátil, Ph.D.**

Ústav elektroniky a měření

Datum zadání diplomové práce:

**25. února 2011**

Termín odevzdání diplomové práce:

**27. května 2011**

Ve Zlíně dne 25. února 2011



prof. Ing. Vladimír Vašek, CSc.

*děkan*



doc. RNDr. Vojtěch Křesálek, CSc.

*ředitel ústavu*

## **ABSTRAKT**

Předložená diplomová práce se zabývá klasifikací výstřelů z ručních zbraní pomocí neuronové sítě. V rámci práce byla vytvořena databáze zvuků, která slouží jako trénovací množina dat pro neuronovou síť. V teoretické části jsou popsány základní principy klasifikace pomocí neuronové sítě, je zde popsána frekvenční analýza signálu, která byla využita pro extrakci charakteristik. V praktické části je navržena a popsána uživatelská aplikace, dále výsledky úspěšnosti klasifikace ručních zbraní do daných kategorií pomocí různě nastavených parametrů neuronové sítě.

Klíčová slova: neuronová síť, klasifikace, dopředná neuronová síť, výstřely ze zbraní

## **ABSTRACT**

This thesis deals the classification of shots from hand gun using neural network. In this thesis was created databases of sounds, that is used as training data set for neural network. The theoretical part is described the basic principles of classification using neural network, there is described frequency signal analysis, which was used for the extraction characteristics. In the practical part of this work is designed and described the user application, and the results of the success of the classification of hand guns into those categories using different set of parameters of neural network.

Keywords: neural network, classification, feedforward, gunshots

**Poděkování:**

Ráda bych touto cestou poděkovala vedoucímu diplomové práce panu Ing. Milanovi Navrátilovi, Ph.D. za odbornou pomoc, připomínky a jeho podporu při vypracovávání této diplomové práce.

Dále děkuji své mamince a celé rodině za jejich pochopení, trpělivost a také za to, že mi umožnili studovat.

**Motto:**

Člověka někdy život nudí, hluboce neuspokojuje, ale na druhé straně si nedokáže představit, co by ho mohlo uspokojivě vylepšit.

*Luke Rhinehart*

**Prohlašuji, že**

- beru na vědomí, že odevzdáním diplomové/bakalářské práce souhlasím se zveřejněním své práce podle zákona č. 111/1998 Sb. o vysokých školách a o změně a doplnění dalších zákonů (zákon o vysokých školách), ve znění pozdějších právních předpisů, bez ohledu na výsledek obhajoby;
- beru na vědomí, že diplomová/bakalářská práce bude uložena v elektronické podobě v univerzitním informačním systému dostupná k prezenčnímu nahlédnutí, že jeden výtisk diplomové/bakalářské práce bude uložen v příruční knihovně Fakulty aplikované informatiky Univerzity Tomáše Bati ve Zlíně a jeden výtisk bude uložen u vedoucího práce;
- byl/a jsem seznámen/a s tím, že na moji diplomovou/bakalářskou práci se plně vztahuje zákon č. 121/2000 Sb. o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon) ve znění pozdějších právních předpisů, zejm. § 35 odst. 3;
- beru na vědomí, že podle § 60 odst. 1 autorského zákona má UTB ve Zlíně právo na uzavření licenční smlouvy o užití školního díla v rozsahu § 12 odst. 4 autorského zákona;
- beru na vědomí, že podle § 60 odst. 2 a 3 autorského zákona mohu užít své dílo – diplomovou/bakalářskou práci nebo poskytnout licenci k jejímu využití jen s předchozím písemným souhlasem Univerzity Tomáše Bati ve Zlíně, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly Univerzitou Tomáše Bati ve Zlíně na vytvoření díla vynaloženy (až do jejich skutečné výše);
- beru na vědomí, že pokud bylo k vypracování diplomové/bakalářské práce využito softwaru poskytnutého Univerzitou Tomáše Bati ve Zlíně nebo jinými subjekty pouze ke studijním a výzkumným účelům (tedy pouze k nekomerčnímu využití), nelze výsledky diplomové/bakalářské práce využít ke komerčním účelům;
- beru na vědomí, že pokud je výstupem diplomové/bakalářské práce jakýkoliv softwarový produkt, považují se za součást práce rovněž i zdrojové kódy, popř. soubory, ze kterých se projekt skládá. Neodevzdání této součásti může být důvodem k neobhájení práce.

**Prohlašuji,**

- že jsem na diplomové práci pracoval samostatně a použitou literaturu jsem citoval. V případě publikace výsledků budu uveden jako spoluautor.
- že odevzdaná verze diplomové práce a verze elektronická nahraná do IS/STAG jsou totožné.

Ve Zlíně

.....  
podpis diplomanta

**OBSAH**

<b>ÚVOD</b> .....	<b>9</b>
<b>I TEORETICKÁ ČÁST</b> .....	<b>11</b>
<b>1 ROZPOZNÁVÁNÍ A KLASIFIKACE</b> .....	<b>12</b>
1.1 ROZPOZNÁVÁNÍ.....	12
1.1.1 Rozpoznávání audio zvuku .....	12
1.2 KLASIFIKACE.....	13
1.3 NEURONOVÉ SÍTĚ .....	13
1.3.1 Základní pojmy neuronových sítí.....	14
1.3.2 Biologický neuron .....	14
1.3.3 Umělý neuron.....	15
1.3.4 Historie .....	17
1.3.5 Učení neuronové sítě.....	18
1.3.6 Rozdělení neuronových sítí.....	18
1.3.7 Algoritmus Back Propagation (BPG).....	20
1.3.8 Topologie neuronových sítí.....	21
1.3.9 Výhody neuronových sítí .....	22
1.3.10 Nevýhody neuronových sítí .....	22
1.4 DALŠÍ METODY KLASIFIKACE .....	23
<b>2 AUDIO SIGNÁL</b> .....	<b>24</b>
2.1 ZÁZNAM AUDIO SIGNÁLŮ .....	24
2.2 ZPRACOVÁNÍ DIGITÁLNÍCH SIGNÁLŮ .....	24
<b>3 EXTRAKCE CHARAKTERISTIK</b> .....	<b>26</b>
3.1 FREKVENČNÍ ANALÝZA .....	28
3.1.1 Fourierova transformace .....	29
3.1.2 Diskrétní Fourierova transformace.....	29
3.2 EXTRAHOVÁNÍ CHARAKTERISTIK .....	30
3.2.1 Výkonové spektrum signálu.....	31
3.2.2 Autokorelační funkce .....	32
3.2.3 Akustický tlak .....	33
3.2.4 Centrální momenty.....	34
<b>II PRAKTICKÁ ČÁST</b> .....	<b>35</b>
<b>4 ÚVOD DO PRAKTICKÉ ČÁSTI</b> .....	<b>36</b>
4.1.1 Popis databází zvuků.....	37
Použité zbraně .....	40
4.2 POUŽITÉ MĚŘICÍ PŘÍSTROJE .....	41
4.2.1 Mikrofonní jednotky .....	41
4.2.2 Jednotka pro sběr dat Advantech USB-4716 .....	43
4.3 POPIS UŽIVATELSKÉ APLIKACE.....	45
4.3.1 Grafické zobrazení aplikace .....	45
4.3.2 Proces trénování .....	46

4.4	ALGORITMUS.....	48
4.5	ÚVOD DO EXPERIMENTÁLNÍ PRÁCE .....	49
4.5.1	Popis tabulky .....	51
4.5.2	Testovací cyklus č.1 .....	53
4.5.3	Testovací cyklus č. 2 .....	56
4.5.4	Testovací cyklus č. 3 .....	59
4.5.5	Testovací cyklus č. 4 .....	62
	<b>ZÁVĚR .....</b>	<b>65</b>
	<b>ZÁVĚR V ANGLIČTINĚ.....</b>	<b>67</b>
	<b>SEZNAM POUŽITÉ LITERATURY.....</b>	<b>69</b>
	<b>SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK.....</b>	<b>72</b>
	<b>SEZNAM OBRÁZKŮ .....</b>	<b>73</b>
	<b>SEZNAM TABULEK.....</b>	<b>74</b>
	<b>SEZNAM PŘÍLOH.....</b>	<b>75</b>

## ÚVOD

V dnešní vyspělé době vědy a techniky není divu, že se lidský mozek a jeho funkce promítají v počítačové technice. Umělé neuronové sítě jsou relativně novou technologií umělé inteligence. První model umělého neuronu byl vytvořen už v roce 1943. Inspirací těchto umělých neuronových sítí je a byla činnost biologických neuronových systémů. Umělá neuronová síť je aplikovatelná na řešení mnoha technických problémů.

Existuje mnoho způsobů klasifikace audio signálů, přičemž tato diplomová práce je zaměřena na klasifikaci audio signálů pomocí umělé neuronové sítě. Tato metoda je v dnešní době velice rozšířená a má velké využití jak při rozpoznávání, tak při klasifikaci audio signálů. Například se využívá při rozpoznávání řeči, kde neuronové sítě pomáhají převádět řeč na text, nebo na klasifikaci hudebních žánrů zde jsou zase neuronové sítě schopny rozlišit, o jaký hudební žánr jde, jestli rock, jazz apod. Pokud opomeneme klasifikaci přímo audio signálu, tak neuronové sítě mají také velké zastoupení v jiných vědních oborech, například v bezpečnostních technologiích napomáhají rozpoznat poznávací značku na automobilech a v mnoha dalších oborech ať už je to zdravotnictví či ekonomika. Jak jde vidět, neuronové sítě mají velké využití nejen v klasifikaci audio signálu.

K tomu, aby neuronová síť byla schopna správně rozpoznávat neznámé výstřely, je potřeba ji vhodně naučit, tzn. poskytnout jí kvalitní trénovací množinu. Za tímto účelem bylo provedeno měření v reálných podmínkách v uzavřené vnitřní střelnici, kde byly použity ruční zbraně. Měření spočívalo v záznamu audio signálů při výstřelech z daných zbraní. Z těchto zvukových projevů byly získány charakteristické rysy (tzv. feature vectory), které představovaly vstupní data pro neuronovou síť.

Teoretická část je zaměřena na problematiku neuronových sítí a jejich funkci. Také na problematiku audio signálu a jeho dalším zpracováním. Praktická část je zaměřena na klasifikaci výstřelů ze zbraní.

Praktickou část představuje nejen měření v terénu ale i úpravu stávajícího programu v prostředí MATLAB 7.9.0 (R2009b). Součástí praktické části, je návrh nastavení počtu skrytých vrstev a v nich takový počet neuronů, aby navržená neuronová síť s co možná největší úspěšností klasifikovala příkládané vzory. Síť bude vícevrstvá a

bude využívat učící algoritmus backpropagation jako způsob učení. Výsledky, které budou na základě této práce zjištěny, budou zpracovány a následně vyhodnoceny.

## **I. TEORETICKÁ ČÁST**

## 1 ROZPOZNÁVÁNÍ A KLASIFIKACE

Mezi rozpoznávání a klasifikací je rozdíl v tom jak zařazujeme objekty a jejich společné vlastnosti. Přičemž objekty mohou být vizuální, akustické a taktilní. Když klasifikujeme objekt, tak ho zařazujeme do předem známého, pevného počtu tříd. U rozpoznávání počet tříd není předem znám a třídy identifikujeme až během vlastního rozpoznávání.

### 1.1 Rozpoznávání

Úloha rozpoznávání spočívá v zařazování objektů, popř. jevů a situací reálného světa, do tříd. Každý z posuzovaných jevů je při dostatečně podrobném popisu jedinečný. Nejprve je třeba určit hledisko, z něhož budou objekty posuzovány, tj. určit veličiny, které je charakterizují. Těmto veličinám je třeba stanovit časoprostorovou rozlišovací úroveň, tedy přesnost a frekvenci, s jakou budou měřeny. Získaná data jsou vstupními údaji pro rozpoznávání. [3]

#### 1.1.1 Rozpoznávání audio zvuku

Zvuk je každé mechanické vlnění v látkovém prostředí, které je schopno vyvolat v lidském uchu sluchový vjem.

##### *Rozpoznávání audio zvuku člověkem*

Zvuky vnímáme sluchovým orgánem, jehož částí jsou ušní boltec, který zachycuje zvukovou energii pak zvukovod, který vede zachycenou zvukovou energii k bláně bubínku a ten se rezonancí rozkmitává a předává dále kmity prostřednictvím kladívka, kovádlíky a třmínku. Eustachova trubice vede ze středního ucha do ústní dutiny a vyrovnává tak případné přetlaky. Oválné okénko je jemná membrána tvořící přístup k vnitřnímu uchu.

Člověk je schopen vnímat zvuky, jejichž frekvence leží v intervalu 20 až 20000 Hz. Jak horní, tak dolní hranice slyšitelnosti jsou individuálně různé a závislé hlavně na věku. Čím je člověk starší, tím více je snížena schopnost vnímat vysoké frekvence.

### *Rozpoznávání audio zvuku počítačem*

Zařízení, které přemění analogový zvuk na elektrické impulzy, se nazývá mikrofon. S těmito elektrickými impulzy se pak dále pracuje, tak abychom dosáhli na zvuk v digitální podobě. Existuje mnoho druhů mikrofonů, liší se tím, jakým způsobem snímají zvuk. Po-té co je zvuk sejmuto mikrofonem se musí převést pomocí A/D převodníku do digitální podoby. Více o digitalizaci v kapitole 2 audio signál

## **1.2 Klasifikace**

Klasifikace je jako taková podoborem teorie rozhodování. Takové tvrzení vychází z předpokladu, že každý pozorovaný vzor musí náležet do určité kategorie nebo třídy, která může být považována za model tohoto vzoru. Bez ohledu na odlišnosti mezi jednotlivými vzory existuje skupina příznaků, které jsou podobné ve vzorech stejné třídy a naopak rozdílné u odlišných tříd. Takové příznaky mohou být použity k rozhodnutí o příslušnosti ke třídě[9]

Použití UNS<sup>1</sup> 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 naučena, začali ji předkládat záznamy, které předtím neviděla. Ukázalo se, že UNS pozná anomálii tak dobře, jako by rozhodovala skupina lékařů[9]

## **1.3 Neuronové sítě**

UNS jsou jednou z oblastí, které se v poslední době věnuje značná pozornost. Příčina jejich úspěchu je přisuzována jejich vlastnostem, které je možno s úspěchem využít v řadě oborů a vědních disciplín.

Cílem této kapitoly je vysvětlení základních pojmů, ukázat rozdíl mezi neuronovými sítěmi s učením s učitelem a s učením bez učitele. A nastínit další rozdělení UNS.

---

<sup>1</sup> UNS - umělé neuronové sítě

Přednostmi neuronových sítí je modelování chování lidského mozku. Lidský mozek obsahuje cca  $10^{11}$  až  $10^{14}$  neuronů uložených v šedé kůře mozkové a synapse jsou realizovány v rozsahu asi  $10^4$  na jeden neuron. Vytvoření úplného umělého mozku se všemi schopnostmi je asi nemožné. Je však možné simulovat alespoň některé funkce lidského myšlení. To jsou potom neuronové sítě. Jejich podstatou je modelování struktury a činnosti biologických neuronových sítí.

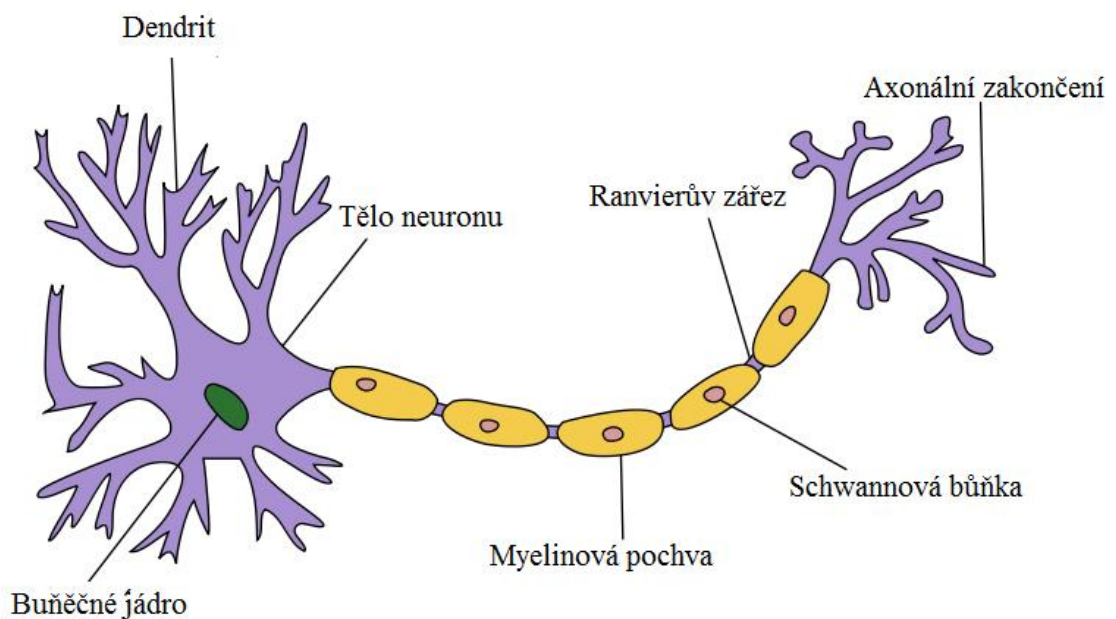
Základními vlastnostmi neuronové sítě je schopnost učit se, tedy nacházet závislosti v trénovacích datech a ty reprezentovat pomocí synaptických vah, a schopnost zevšeobecňování získaných znalostí. Zevšeobecňování je schopnost správně reagovat také na neznámé vstupy, na které nebyla neuronová síť naučena.[10]

### 1.3.1 Základní pojmy neuronových sítí

Neuronová síť se skládá z umělých neuronů, které vznikly na základě biologického neuronu. Pro lepší pochopení neuronových sítí si první vysvětlíme funkci biologického neuronu.

### 1.3.2 Biologický neuron

Je to buňka schopná přijmout, vést a zpracovat a odpovědět na signály. Přenáší a zpracovává informace z vnitřního a vnějšího prostředí a tím podmiňuje schopnost lidského organismu na ně reagovat.

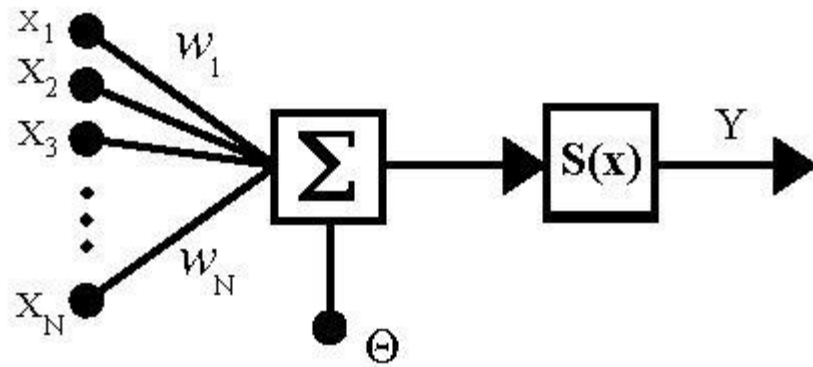


Obrázek 1: Biologický neuron

Neuron má svou buněčnou membránu, která ho odděluje od zevního prostředí, ohraničuje jeho tělo a chrání ho před vnějšími vlivy. Z těla neuronu vystupují četné, bohatě rozvětvené výběžky – dendrity a axony. Dendrity vedou vzruch směrem k buňce a tvoří vstupy do neuronu. Axon je v každém neuronu jen jeden a tvoří výstup z neuronu. Konce axonu jsou opatřeny tzv. synapsemi. Jsou to velmi složité útvary, které zajišťují přenos informací mezi navzájem spolupracujícími neurony – vedou informaci k dalším neuronům. Spojením neuronů pomocí synapsí vzniká biologická neuronová síť. K aktivaci neuronů dochází v případě, že souhrn vstupních signálů překročí určitou prahovou hodnotu. Následně neuron vyšle do svého axonu výstupní signál jako reakci na dané vstupy a informace se tak šíří neuronovou sítí. Pokud je však souhrn vstupních podnětů menší než je prahová hodnota, žádná reakce v neuronu nevznikne a na výstupu z neuronu je signál, odpovídající pasivnímu stavu neuronu. Tok signálu neuronovou sítí probíhá vždy pouze z axonu jedné buňky do dendritů buněk dalších. Základním prvkem neuronových sítí je neuron. V porovnání s biologickým neuronem lze pomocí počítačů simulovat mnohem rychlejší neuron než je ten biologický. Problémem je však velké množství biologických neuronů a množství synapsí v mozku. Takto masivně paralelní neuronovou síť není ani v blízké budoucnosti simulovat. Síla lidského mozku je tedy ve využití velkého množství pomalých neuronů, seskupených do velmi složité masivně paralelní sítě, jejíž velikost a struktura není v simulačních procesech v současnosti ani zdaleka napodobitelná.

### 1.3.3 Umělý neuron

Základní jednotka umělých neuronových sítí. Je to samostatně pracující element s mnoha vstupy a většinou jedním výstupem. Výstupní signál je závislý pouze na vstupních úrovních. Princip neuronu spočívá v procesu učení, při kterém se celý systém adaptuje podle optimalizačních algoritmů, aby co nejlépe vyřešil zadanou úlohu. [10]



Obrázek 2: Model umělého neuronu Zdroj: [10]

$x=[x_1, x_2, \dots, x_n]$  – vstupy neuronu (výstupy z předcházející vrstvy),  $i=1, 2, \dots, n$ ,

$n$  – počet vstupů (počet neuronů v předcházející vrstvě),

$w=[w_1, w_2, \dots, w_n]$  - jsou vektory vah neuronu,

$\theta$  – je práh neuronu,

$\Sigma$  – vstupní potenciál neuronu,

$S(x)$  – výstupní funkce neuronu,

$x$  – aktivační funkce neuronu,

$y$  – výstup neuronu.

$$y = S\left(\sum_{i=1}^N w_i x_i + \theta\right) = S(W^T X + \theta), \quad (1)$$

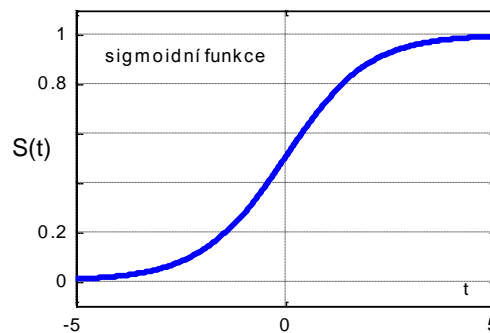
Vstupy  $x_1, x_2, \dots, x_n$  modelují dendridy, synaptické váhy  $w_1, w_2, \dots, w_n$  modelují synapse. Výstup  $y$  simuluje činnost axonu. Každý neuron obsahuje konečný počet vstupů  $x_n$  a jediný výstup  $y$ . V každém neuronu se vstupní hodnoty transformují na výstup pomocí minimálně dvou výpočetních procedur, vypočtu vstupního potenciálu  $\Sigma$  a aktivační funkce  $S$ . Na neuronovou síť je třeba nahlížet jako na dynamický systém, tj. systém, jehož vlastnosti se mění v závislosti na čase. [9]

Výběr aktivační funkce zaleží na tom, co chceme pomocí neuronové sítě řešit. Pokud řešíme otázky typu ano/ne stačí binární funkce. Pokud už řešíme otázky identifikace, klasifikace apod. jsou k tomu potřeba jiné funkce viz dále.

Aktivační funkce, která je využita v této práci:

$$S(t) = \frac{1}{1 + e^{-t}} \quad (2)$$

Funkce odvozená od přenosové funkce biologického neuronu. Tato funkce se nazývá sigmoidní funkce. Poukazuje na ni níže uvedený obrázek.



Obrázek 3: Sigmoidní funkce

Pro přehlednost uvádím další aktivační funkce, které se používají:

- *Lineární*
- *Skoková*
- *Hyperbolický tangens*

### 1.3.4 Historie

Za počátek historie neuronových sítí se považuje rok 1943, kdy Warrena McCulloch a Walter Pittse vytvořili první model neuronu. V roce 1949 vydal Donald Hebb knihu *Organization of Behavior*, ve které bylo navrženo učící pravidlo samostatného neuronu (dnes známé jako Hebbovské učení). Velkou smůlou pro neuronové sítě byla diskreditační kampaň vedená Martinem Minským (mimo jiné konstruktér prvního neuropočítače) a Seymourem Papertem. Tato kampaň vyvrcholila roku 1969, kdy byl knižně vydán článek *Perceptron*, který zdůrazňoval neschopnost neuronu naučit se řešit problém XOR<sup>2</sup>. Až do roku 1982 se v této oblasti nedělo nic mimořádného a nebyly ani vypisovány granty. Jedině za podmínky mezioborovosti. V této době vydal fyzik John

---

<sup>2</sup> XOR - V logice a matematice je exkluzivní disjunkce označením pro „buď ..., anebo ...“. Například „Počítač je buď zapnutý anebo vypnutý.“ je exkluzivní disjunkce.

Hopfield dva články. Ale teprve až roku 1986 byl publikován článek o učícím algoritmu Back-Propagation. Tento algoritmus byl určen pro učení vrstevnatých neuronových sítí. Byl "objeven" vícekrát, protože předtím byl publikován v letech 1969, 1974, 1985, ale pokaždé zanikl.

Velkým úspěchem byl systém NETtalk, který umožňoval převod anglického psaného textu do mluvené angličtiny a dosahoval lepších výsledků než DECtalk, na jehož vývoji se strávilo několik let - převážně formulováním složitých pravidel.[11]

### 1.3.5 Učení neuronové sítě

Aby nám neuronová síť vykonala to, co po ní vyžadujeme, musíme jí to naučit. Učení lze charakterizovat jako získávání znalostí o světě, včetně pocitů, reakcí a stavů. Celkový život neuronových sítí lze rozdělit do dvou fází:

- Fáze učení, kdy se znalosti ukládají do synaptických vah neuronové sítě. Jestliže  $W$  je matice všech synaptických vah neuronové sítě, pak pod učením se rozumí stav, kdy platí

$$\frac{\partial W}{\partial t} \neq 0. \quad (3)$$

Synaptické váhy se v průběhu učení mění. Pojem učení u neuronové sítě je synonymem pojmu adaptace neuronové sítě. Jde tedy o sbírání znalostí, resp. Jejich uchovávání. Učení je proces, ve kterém se parametry neuronové sítě (synaptické váhy) mění na základě nějakých pravidel determinovaných typem učení neuronové sítě. Pod učením se rozumí adaptace neuronové sítě, která po ukončení učení bude nositelkou znalostí získaných v průběhu učení.

- Fáze života, kdy se získané znalosti využívají ve prospěch řešení nějakého problému. Synaptické váhy se tedy v této fázi nemění. Jde o stav kdy

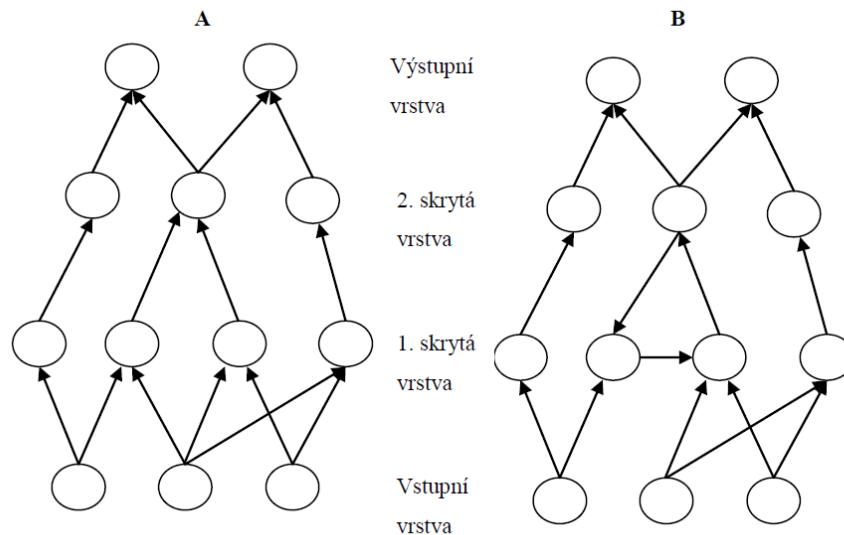
$$\frac{\partial W}{\partial t} = 0. \quad (4)$$

### 1.3.6 Rozdělení neuronových sítí

- Podle metody šíření informací

Rozlišují se dva základní typy struktur neuronové sítě, dopředná a rekurentní.

**Rekurentní (RCNN,RNN)**, (obr.: 3B) – zde existuje skupina neuronů, která je zapojena v kruhu. Synapse se orientují různými směry. V takové struktuře může být neuron současně vstupní i výstupní (tzv.: duální neurony). Rekurentní neuronové sítě představují dynamicky složitější systém než dopředné neuronové sítě. I zde lze sítě rozdělit na rekurentní sítě s učitelem nebo rekurentní sítě bez učitele.



Obrázek 4: Neuronové sítě. A - dopředná síť

B – rekurentní síť. Zdroj [1]

Nejjednodušším typem rekurentních neuronových sítí s učitelem jsou Hopfieldovy sítě. Tato síť se skládá z duálních neuronů. Vstupem do jednotlivých neuronů je vstup z reálného světa a vstup z ostatních neuronů.

**Dopředné sítě (FFNN)**, (obr.: 3A) – zde se signál šíří po synapsích jedním směrem vpřed. Od vstupu směrem k výstupu. Nevytváří žádný uzavřený cyklus.

V této práci je použita dopředná neuronová síť. Proto si tuto problematiku vysvětlíme trochu podrobněji.

Dopředné neuronové sítě jsou nejznámějším typem neuronových sítí. Vyznačují se postupným zpracováním informace od vstupu směrem k výstupu. Neurony každé z vrstev jsou propojeny s neurony vrstvy předchozí a jejich výstupy jsou dále vedeny do vstupů neuronů vrstvy následující. Síť tedy neobsahuje zpětné vazby a je ji možno popsat algebraickými rovnicemi.[17]

Pro učení dopředné sítě o více vrstvách se používá algoritmus Back Propagation. Jeho objevení v roce 1986 vědci Rumelhart a McClelland umožnilo vyřešit problém XOR, který byl pro neuronové sítě do té doby neřešitelným a byl nemalou příčinou k zastavení výzkumu UNS v tehdejší době [13]. V dnešní době je algoritmus Back Propagation nejpoužívanější metodou učení na bázi učení s učitelem

### 1.3.7 Algoritmus Back Propagation (BPG)

Neboli algoritmus zpětného šíření chyby. Back Propagation je metoda výuky umělých neuronových sítí. Tato metoda je použitelná pouze u sítí s dopředným šířením.

*Učení podle tohoto algoritmu probíhá ve 3 fázích:*

1. Na vstup neuronové sítě je přiváděn vektor resp. pro více vzorů matice vstupních parametrů. V našem případě jsou to vzorky z databáze signálů.
2. Po průchodu neuronovou sítí je výsledek porovnán s požadovanou hodnotou, je spočítána chyba, ta se zpětně přepočítává do předchozích vrstev a synaptické váhy představující paměť jsou opraveny.
3. Do opravené sítě je znovu přiveden vstupní vektor resp. matice a proces se opakuje. Jedná se o iterativní proces. Hledáme minimum chyby mezi skutečnou (výstupní) hodnotou a požadovanou hodnotou pro všechny vzory učení.

Tato metoda je v praxi nejčastěji používaná a je vhodná pro mnoho aplikací (pokud známe výsledek, ke kterému chceme dojít), např. pro aproximace funkcí, predikci nějakých parametrů nebo jak je to v našem případě pro klasifikaci. Její nevýhodou je velká citlivost na relevantnost vstupních dat a na inicializaci synaptických vah.[18]

- Podle počtu vrstev
  - **Jednovrstvá (Hopfieldova síť)**
  - **Vícevrstvá (ART síť)**

- Podle stylu učení

Jak se získávají váhy jednotlivých neuronů. Když jsou váhy určeny výpočtovým algoritmem například algoritmus Back Propagation, jde o deterministický styl učení. Více o tomto algoritmu v kapitole 1.3.7.

Jestliže jsou váhy získány pomocí generátoru náhodných čísel, jde o stochastický styl učení.

- **Deterministicky**
- **Stochasticky**
- Podle způsobu učení

Učení sítě obstarává učicí algoritmus, který během učení mění váhy jednotlivých spojů. Teorie učení neuronových sítí je postavená kolem Hebbova zákona, který je inspirován způsobem učení mozku: „Pokud jsou dva neurony v jednom okamžiku aktivní, zesil jejich vazbu a v opačném případě zeslab.“ Naučení sítě je tedy uloženo ve vahách spojů mezi neurony, samotná topologie při se při učení nemění. U učení rozlišujeme dva druhy učení.

- **S učitelem** – máme k dispozici páry vstupní a k němu příslušející výstupní vektor z reálně naměřených dat v našem případě výstřely ze zbraní. Algoritmus porovná požadovaný výstup se skutečným výstupem sítě a upraví váhy spojů tak, aby se výstup přiblížil požadované hodnotě. To provede s každým prvkem trénovací množiny.

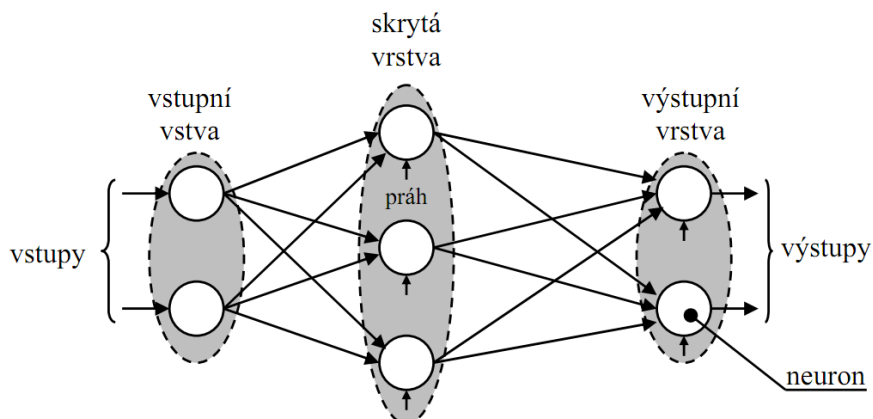
S učením je také spojen jeden z problémů umělých neuronových sítí, a to konkrétně přeučení – síť sice perfektně odpovídá na vstupy z trénovací množiny, ale není schopna generalizace a odpovědi na reálná data jsou nesmyslné.

- **Bez učitele** – zde má síť k dispozici pouze vstupní data a hledá v nich shluky = podobná data. Algoritmus pak posouvá jednotlivé neurony ve stavovém prostoru. Proces aplikace naučených poznatků se nazývá vybavování.[16]

### 1.3.8 Topologie neuronových sítí

Topologie sítě určuje způsob vzájemného propojení neuronů. Většina sítí má neurony organizovány ve vrstvách, které jsou mezi sebou propojeny systémem každý s každým. Každá síť obsahuje vstupní a výstupní vrstvu a může obsahovat několik skrytých vrstev. Podle topologie můžeme sítě v základě rozdělit na přímé a rekurentní. Vstupní/výstupní data označujeme jako vstupní/výstupní vektor, dimenze vektorů odpovídají počtu vstupů/výstupů sítě v určitých případech se mohou shodovat.

V této práci budou použity 2 skryté vrstvy.



Obrázek 5: Příklad dopředné neuronové sítě s jednou skrytou vrstvou [17]

V této práci jsou jako vstupní data použity signály z databáze vzorů. Tyto vstupy jsou dané. Za to další skrytá vrstva už může obsahovat libovolný počet neuronů, tyto neurony si volím v programu. Více nebo méně jak 2 skryté vrstvy nebudou použity. Jako poslední je výstupní vrstva, která je také daná a to podle klasifikačních složek.

### 1.3.9 Výhody neuronových sítí

Typicky se neuronové sítě dají použít tam, kde klasické počítače selhávají. Tam, kde jsou k dispozici příkladová data, která dostatečně pokrývají problémovou oblast. Hlavní výhodou je jejich schopnost učit se.

- *Paralelní zpracování informací*
- *Učící schopnost*
- *Možnost adaptace na změnu parametrů*
- *Odolnost vůči chybám*

### 1.3.10 Nevýhody neuronových sítí

Jako každý program i neuronové sítě mají své nevýhody.

- *Časová náročnost*
- *Sestavení vstupních a výstupních dat*
- *Složitost programování*
- *Nalezení optimálního nastavení pro danou úlohu*

## 1.4 Další metody klasifikace

V této práci jako klasifikační metoda je použita metoda umělých neuronových sítí. Neuronové sítě jsem zvolila proto, že jsou schopny adaptace na změnu parametrů. Jsou vhodné pro úlohy identifikace, aproximace, klasifikace a třídění vzorů. Proto jsou neuronové sítě v předchozí kapitole detailně popsány. Zbylé klasifikační metody uvádím jen pro informaci.

- *Klasifikátor s diskriminační funkcí*
- *Kritérium minimální chyby - Bayesův klasifikátor*
- *Kritérium minimální vzdálenosti*
- *k-NN nejbližší soused*

Toto je stručný přehled nepoužívanějších klasifikačních technik.

## 2 AUDIO SIGNÁL

Signály, jako je řeč a hudba, tvoří základ mezilidské komunikace. V současné době dochází k progresivnímu rozvoji akustické komunikace mezi člověkem a automaty. Akustické signály jsou:

- *Technického původu* – signály uměle generované. Například měřicí signály, ale také i nežádoucí produkty vznikající používáním technických prostředků jako je šum.
- *Přírodního původu*[12]

Signály jsou obvykle zaznamenávány pro jejich pozdější zpracování a analýzu. Upravený audio signál může být poté znovu reprodukován. Zpracování signálu lze v zásadě provádět analogovým, číslicovým, nebo kombinovaným způsobem. Čistě analogový způsob zpracování se již pomalu přesouvá do historie. Základem se stává číslicové (digitální) zpracování audio signálu.

Číslicové zpracování vyžaduje převod analogového signálu na jeho číslicovou reprezentaci. S tím souvisí i zpětný převod digitálního signálu na signál analogový.[12]

### 2.1 Záznam audio signálů

Sběr akustických signálů se uskutečňuje pomocí elektroakustických měničů, které pracují jako snímače signálu. Pro snímání akustických signálů se používají měřicí kondenzátorové mikrofony, elektretové mikrofony, ve vodním prostředí hydrofony, apod. Snímané signály pro přesná měření jsou zesíleny zaznamenávány v analogové nebo číslicové podobě za účelem dalšího zpracování.

### 2.2 Zpracování digitálních signálů

Zpracování digitálního signálu se dnes používá snad v každém odvětví. Jako nejnámější můžeme zařadit zpracování audio a video signálu. Digitální signály jsou zpracovávány nejrůznějšími metodami. Zde jsou uvedeny nejpoužívanější a nejnámější metody zpracování digitálního signálu.

### VZORKOVÁNÍ A KVANTOVÁNÍ

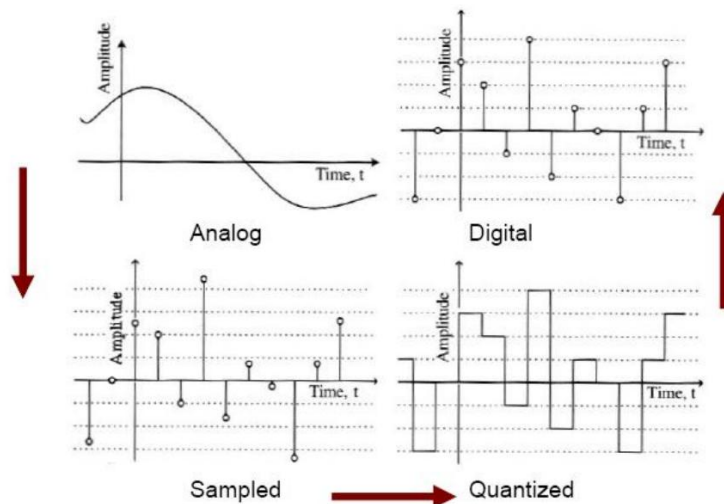
Pro další zpracování analogového signálu v počítači je potřebná jeho digitalizace, která je prováděna na analogově-digitálních převodnicích. V těchto převodnicích probíhají 2 kroky. První je vzorkování a druhý je kvantování. Vzorkování signálu však podléhá

důležitému pravidlu, kterým je Nyquist-Shannonův vzorkovací teorém. Ten říká, že signál se musí vzorkovat minimálně dvojnásobnou frekvencí obsaženou ve vzorkovaném signálu.[22] k případné správné rekonstrukci signálu

Minimální vzorkovací frekvence:

$$f_s \geq 2 \times f_{max} \quad (5)$$

Následující obrázek zobrazuje převod zvuku z analogové podoby do digitální podoby. První obrázek (Analog) obsahuje signál v analogové podobě, který má být upraven do digitální (číslicové) podoby. Dalším krokem je vzorkování signálu (Sampled) – rozdělení a odebrání vzorů. Dále pokračuje kvantování (Quantized). Jako poslední obrázek je už signál v digitální podobě.



Obrázek 6: Digitalizace analogových dat [30]

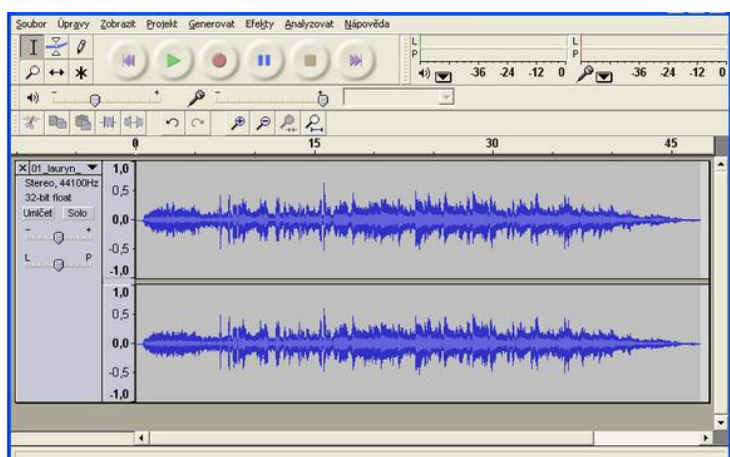
Vzorkování i kvantování nám zajišťuje A/D převodník, který převádí analogový signál na digitální. Na reprodukci je zase potřebný D/A převodník s obrácenou funkcí. Tyto dva převodníky se často používají v jednom čipu, který se označuje kodek (kóduje a dekóduje zvuk z jedné formy druhé).

### 3 EXTRAKCE CHARAKTERISTIK

První fází procesu při zpracování zvuku jeho snímání pomocí mikrofonních jednotek. Druhou fází je zpracování zvuku, jeho záznam a převod do digitální podoby. To vše bylo probráno v předchozí kapitole. Teď je na řadě jeho editace tzn. vystřihování, kopírování apod. V našem případě výstřely ze zbraní musíme dále upravit, tak aby byly použitelné jako trénovací množina vzorů do neuronové sítě. Úprava spočívá ve čtyřech krocích, prvním je právě vystřihování z originálně snímaného signálu. Dalším krokem je přepočítání na akustický tlak. Třetím krokem úpravy je normalizace, která nedělá nic jiného než, že přepočítá pomocí vzorce všechny zaznamenané vzory na jednotnou vzdálenost. Posledním krokem je frekvenční nebo časová analýza, která extrahuje dané příznaky ze signálu. Všechny uvedené výše popsané úpravy jsou dále popsány.

#### a. Snímaný signál

Ve volně šiřitelném programu Audacity na střih byl signál upraven. Náhled do programu je vidět na obrázku č.7.



*Obrázek 7: Signál  
zaznamenaný mikrofonními  
jednotkou*

K tomu abychom mohli dále získané vzorky používat, musíme odstranit případné nežádoucí efekty na datech. V tomto případě se aplikuje vyřezávání. U těchto výstřelů to bylo ořezání délky signálu ze 1.4 s na 0.82 s.

#### b. Úprava signálu

O zvuku je možné hovořit jako o změně tlaku v prostředí. Při výstřelu dochází tedy k změně tlaku vzduchu v čase. Pokud je signál dostatečně ořezán a redukován, tak po

konverzi vzorkovací frekvence na hodnotu 20000 Hz jsou jednotlivé zvukové záznamy přepočítány na akustický tlak pomocí vztahu:

$$y_{pa}(k) = y(k) \cdot \frac{U_r}{C_{mic}}, \quad (6)$$

Kde  $y(k)$  je zaznamenaný zvukový signál ve formátu s pohyblivou řádovou čárkou v rozsahu  $\langle -1; 1 \rangle$ ,  $U_r$  je nastavený napěťový rozsah jednotky pro sběr dat při záznamu vzorů a  $C_{mic}$  je citlivost mikrofону udaná jeho výrobcem (v našem případě  $31,6 \cdot 10^{-3} \text{ V.Pa}^{-1}$ ).

Normalizace slouží pro úpravu dat tak, aby nezáleželo na vzdálenosti mikrofónu a zdroje zvuku.

### c. Normování

Jelikož zaznamenané signály byly naměřeny v různých vzdálenostech, musí se získaná data normovat a to na vzdálenost 10m.

Aby byly všechny vzorky v databázi porovnatelné z hlediska velikosti akustického tlaku, jsou všechny normalizovány na vzdálenost 10m od zdroje zvuku. Pro potlačení vlivu přídavného útlumu prostředí, který je značně závislý na podmínkách šíření zvukové vlny od zdroje k mikrofónu, jsou tímto způsobem přednostně zpracovány vzorky zaznamenané z menších vzdáleností. Pro vlastní přepočet akustického tlaku se použije vztah:

$$y_{pa10m}(k) = y_{pa}(k) \cdot \frac{l_{mic}}{l_{norm}} \quad (7)$$

kde  $l_{mic}$  je vzdálenost mikrofónu od zdroje zvuku při pořizování záznamu a  $l_{norm}$  je nová normovaná vzdálenost pro přepočet akustického tlaku.

Z důvodu přepočtu vzorů na akustický tlak v normované vzdálenosti 10m již není možno vzorky uložit do databáze ve formě standardních formátů WAV<sup>3</sup> (WAVE) souborů. Pro

---

<sup>3</sup> Zvukový formát WAVE vznikl z Microsoft RIFF formátu pro ukládání multimediálních souborů.

uložení byly zvoleny datové soubory v obecném textovém formátu, které lze jednoduše importovat pro další zpracování prakticky do libovolných programových prostředí.

Pokud už je signál upraven tzn. přepočítán na akustický tlak, normován na jednu vzdálenost, můžeme aplikovat další část úpravy vzorů. Tato úprava spočívá v aplikaci časové či frekvenční analýzy.

#### **d. Analýza signálu**

Analýzu signálů můžeme rozdělit podle toho, zda analýza probíhá v časové nebo kmitočtové oblasti. Pod pojmem analýzy signálů v časové oblasti rozumíme statickou analýzu, výpočet a analýzu korelačních funkcí, konvolučních funkcí, koherence apod. Analýzu v kmitočtové oblasti představuje například kmitočtová analýza akustických signálů pomocí filtrů s konstantní nebo proměnnou šíří pásma. Dále pak číslicová analýza signálů, která se provádí pomocí číslicové filtrace, nebo např. pomocí algoritmů pro výpočet diskrétní Fourierovy transformace. [12]

Základní úlohou analýzy signálu je popsat vlastnosti signálu jeho souhrnnými charakteristikami.[5]

V případě této diplomové práce je použita analýza jak v kmitočtové oblasti, tak v časové oblasti.

Frekvenční oblast:

- *Průměrování spektra*
- *Korelační funkce*

Z časové oblasti jsou použity tyto analýzy:

- *Akustický tlak*
- *Centrální momenty*

### **3.1 Frekvenční analýza**

Pro transformaci signálu z časové nebo prostorové oblasti do frekvenčního pásma můžeme využít běžnější Fourierovu transformaci. Fourierova transformace nám říká informace o frekvenční charakteristice signálu – převádí signál do komponent fáze a amplitudy. Účelem frekvenční analýzy je odhalit, které frekvence jsou v signálu obsaženy a které naopak chybí. [24]

Všechny analyzované audio signály byly porovnány se vzory v databázi pomocí korelační analýzy nebo spektra. Výsledkem porovnání je koeficient v rozsahu 0 až 1 určující podobnost se vzorkem v databázi. Vzorek, jehož korelační koeficient se nejvíce blíží k 1, se potom vyhodnotí jako nejvíce podobný analyzovanému vzorku a přiřadí se mu příslušná klasifikační třída.

### 3.1.1 Fourierova transformace

Fourierova transformace je známá od první poloviny minulého století. Již více než sto let se, s mírně kolísající intenzitou používá k popisu systému a při řešení různých problémů v matematice, fyzice a v mnoha jiných oborech. K nesporným výhodám transformace je přechod od časových závislostí k závislostem frekvenčním. Fourierova transformace se používá pro řešení mnoha úkolů v mnoha oborech.

Fourierova transformace (FT) je z matematického pohledu operace, která provádí transformaci funkce v časové oblasti na funkci ve frekvenční oblasti. Fourierova transformace vychází z předpokladu, že každý signál lze vyjádřit jako superpozici nekonečně mnoha sinusových signálů. Výsledkem Fourierovy transformace jsou Fourierovy koeficienty, které vyjadřují, jak daný sinusový průběh přispívá k celkovému signálu. Analyzované signály lze rozdělit na spojité a diskrétní a ty mohou být dále periodické a aperiodické. Podle typu signálu je nutno použít správný typ Fourierovy transformace. Jelikož pracujeme s diskrétním signálem je v této práci použita diskrétní Fourierova transformace. [12]

### 3.1.2 Diskrétní Fourierova transformace

V případě číslicového zpracování signálu, kdy pracujeme s diskrétní reprezentací signálu  $x[n]$  s délkou  $N$  vzorů, přecházíme při zpracování signálu na diskrétní Fourierovu transformaci (DTF), která je vyjádřena vztahem:

$$X(k) = \frac{1}{N} \sum_{n=0}^{N-1} x[n] e^{-j \frac{2\pi kn}{N}}, \quad k = 0, 1, \dots, N-1 \quad (8)$$

Tvar zpětné diskrétní Fourierovy transformace (IDFT) je

$$x(n) = \sum_{k=0}^{N-1} X[k] e^{j \frac{2\pi kn}{N}} \quad (9)$$

,  $n=0,1, \dots, N-1$ .

Jediným typem Fourierovy transformace, který může být použit pro digitální zpracování signálů je právě diskrétní Fourierova transformace, protože dnešní digitální technika dokáže pracovat pouze s diskrétními signály o konečné délce[12]

### 3.2 Výběr charakteristik

Dalším krokem při zpracování audio signálu pro vstup do umělé neuronové sítě je výběr vhodných charakteristik z audio signálu. Díky těmto charakteristikám bude stále možné neznámý audio signál klasifikovat, avšak tímto krokem zredukujeme množství informací přiváděné na vstup neuronové sítě. Na extrakci charakteristických rysů byly použity následující analýzy.

### 3.2.1 Výkonové spektrum signálu

Výkonové spektrum signálu udává rozložení výkonu signálu podél frekvenční osy. Spektrum signálu je určeno jeho Fourierovou transformací. Na získání charakteristického rysu ze signálu nám nestačí vypočítat jen Fourierovu transformaci, jelikož signál je stále velmi obsáhlý. Proto použijeme průměrování.

Proces průměrování zahrne několik jednotlivých časových záznamů neboli vzorů. Průměrujeme  $m$  vzorů, každý o nějaké délce  $t$ .

U průměrování spektra je možnost právě nastavení průměrovaného kroku (kolik  $m$  vzorů má být zprůměrováno). Čím více  $m$  vzorů je průměrováno, tím je signál hladší. Na obrázku 8 je vidět kolik bodů je na originálním vstupu a na výstupu po průměrování.

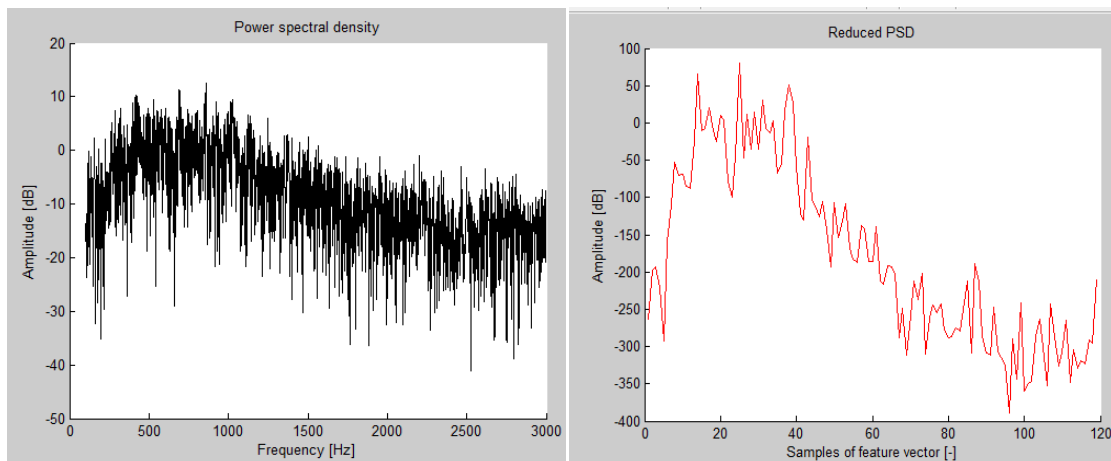
Grafy jsou zobrazeny za těchto podmínek:

Originální signál:

- Body k frekvenční analýze: 16 384 bodů
- Min. frekvence: 100 Hz
- Max. frekvence: 3000 Hz

Redukovaný signál:

- Krok průměrování: 10



Obrázek 8: Originální signál A - Průměrovaný signál B

Z těchto grafů lze vyčíst, jaký vliv má průměrovací krok na výslednou redukci signál.

### 3.2.2 Autokorelační funkce

Další charakteristikou, která nám bude nápomocná v extrakci charakteristických rysů ze signálu ve frekvenční oblasti je autokorelační funkce. Při použití korelačních funkcí je potlačen šum, neboť korelační funkce porovnává, jak moc se podobají (míru podobnosti) signály posunuté navzájem o  $\tau$  (časový úsek). Autokorelace je definována jako korelace napříč signálem v proměnných intervalech času či prostoru.[22]

Tato analýza v programu využívá hned dvě redukce signálu a to redukční faktor a klouzavý průměr. Redukční faktor zredukuje počet bodů v audio signálu a klouzavý průměr tento zredukováný signál ještě vyhladí. Obr. 9 ukazuje, jak vypadá signál po použití právě těchto dvou redukčních kroků.

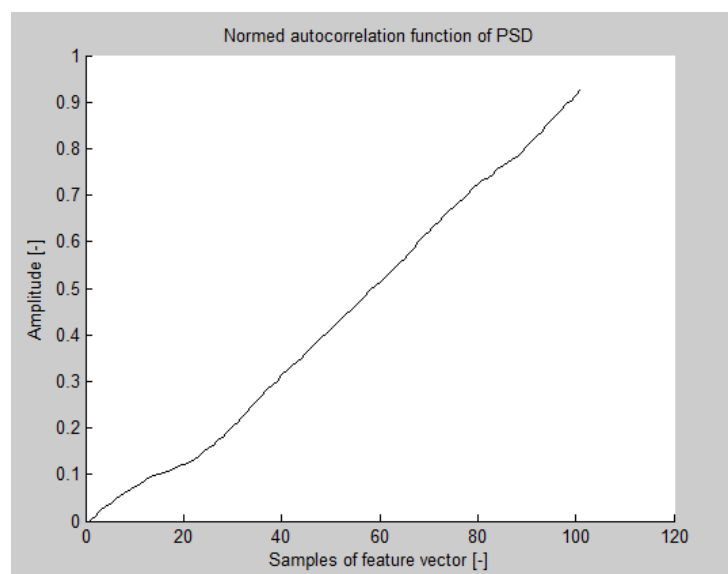
Graf je zobrazen za těchto podmínek:

Originální signál:

- Body k frekvenční analýze: 16 384 bodů
- Min. frekvence: 100 Hz
- Max. frekvence: 5000 Hz

Redukovaný signál:

- Redukční faktor: 40
- Klouzavý průměr: 17



Obrázek 9: Redukovaný signál pomocí autokorelační funkce

### 3.2.3 Akustický tlak

Akustický tlak, je následkem změn tlaku vzduchu, způsobených zvukovými vlnami. Nejnižší akustický tlak, který je lidským uchem vnímán, se nazývá práh slyšitelnosti. Nejvyšší akustický tlak, který ještě lidské ucho snese, se nazývá práh bolesti[28]

Akustický tlak patří do časové analýzy, tzn., že tlak je závislý na čase. Proto přímo v programu je možnost tento čas ořezat a i tím redukovat signál. Další možností jak redukovat signál je redukční faktor. Tento faktor je volen na základě velké hustoty signálu, jak jde vidět na obrázku A. redukční faktor je u této charakteristiky nastavitelný v samotném programu. Tento faktor odebere ze signálu tolik bodů kolik je nastaveno. Tyto dva kroky umožní zredukování signálu, který bude sloužit jako vstupní množina do umělé neuronové sítě. Na obrázku 10B(Redukovaný akustický tlak), lze vidět jaký vliv má redukční faktor a časové zkrácení na výsledný signál.

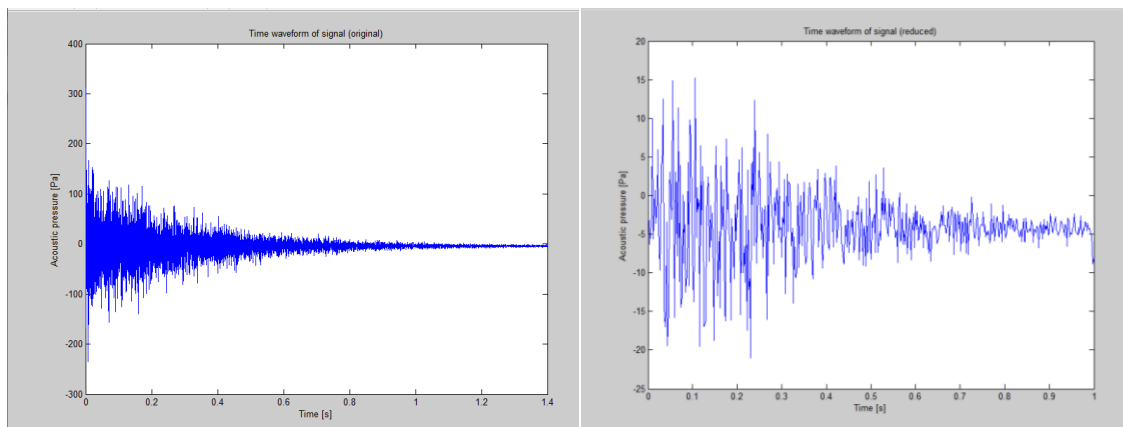
Grafy jsou pořízeny z programu MATLAB.

#### Originální signál:

- Délka vzorku  $t=1.4s$
- Frekvence: 20 000 Hz

#### Redukovaný signál:

- Délka vzorku  $t=1.0 s$
- Posun signálu: 0.1 s
- Redukční faktor: 40



Obrázek 10: Originální signál A - Redukovaný akustický tlak B

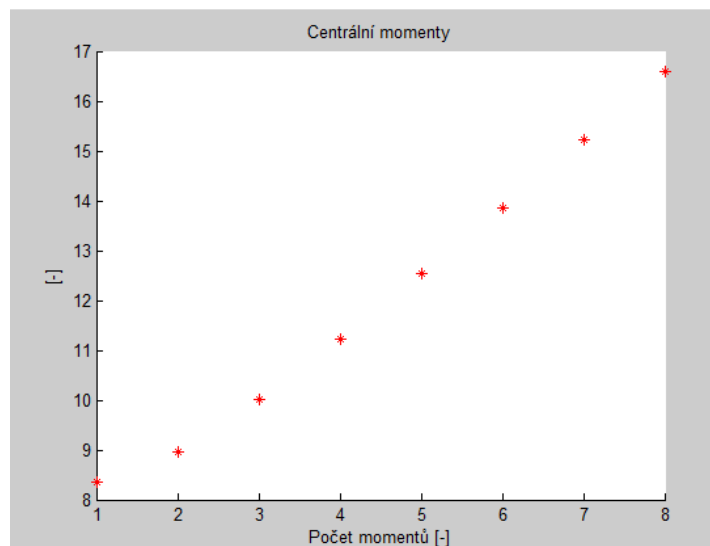
### 3.2.4 Centrální momenty

Centrální momenty jsou charakteristikami náhodných veličin. Spektrální moment je těžiště nebo také centrum spektra. Náhodné veličiny přiřazují reálné číslo. První moment je střední hodnota.

Graf je zobrazen za těchto podmínek:

Redukovaný signál:

- Počet centrálních momentů: 8



Obrázek 11: Centrální momenty

Tato kapitola stručně popsala charakteristiky jak z frekvenční oblasti tak z časové, které v praktické části budou extrahovat charakteristické rysy ze vstupních signálů. Tyto signály už do neuronové sítě budou vstupovat redukované a přichystané ke konečné klasifikaci. V praktické části bude zjištěno, která z analýz ať už frekvenčních či časových nejlépe zachová informaci o signálu a bude vykazovat spolu s neuronovou sítí největší úspěšnost při klasifikaci.

## **II. PRAKTICKÁ ČÁST**

## 4 ÚVOD DO PRAKTICKÉ ČÁSTI

Aby se neuronová síť správně naučila a zvuky správně klasifikovala, je nutné jí poskytnout kvalitní učící množinu zvuků. Proto praktická část diplomové práce spočívala ve vytvoření vzorové množiny výstřelů z různých zbraní ve vzdálenostech 10,15,25 m. Dále vytvoření algoritmu, který bude tuto množinu zvuků klasifikovat. Střelba probíhala v areálu TOMA a.s v Otrokovcích ve střelnici Guncenter.



*Obrázek 12: Guncenter Otrokovice*



*Obrázek 13: Tvorba databáze*

Jelikož mám zbrojní průkaz a provozovatelé v Guncentru nám vyšli vstříc s našimi požadavky jako např. rozmístění mikrofonů ve střelišti nebo půjčení většího množství zbraní, tak nebyl problém si tuto databázi výstřelů vytvořit.

V praktické části popíšu, jak vznikaly databáze výstřelů, kde v tabulkách bude zřejmé, co je přímo jejich obsahem. Budu se zde věnovat i použitým měřicím přístrojům, které zaznamenávaly výstřely ve střelišti, které následně tyto signály digitalizovaly.

Digitalizovaná data se následně normovaly. Na normování všech vstupních signálů jsme použili akustický tlak, takže všechny vzorky byly přepočítány vzorcem na akustický tlak. O tomto normování více v teoretické části v kapitole 3 Extrahování charakteristik.

Další kapitola bude obsahovat popis samotného programu. Jak vznikal, jak funguje a jeho ovládání. Vše bude dokumentováno obrázky přímo z programového prostředí Matlab.

Úspěšnost trénování neuronové sítě jsou závislé na počtu neuronů ve skrytých vrstvách, ale také nastavení parametrů při analýze signálu. Nejlepší propojení se budu snažit najít a to experimentální činností.

#### **4.1.1 Popis databází zvuků**

Databáze je tvořena ze sedmnácti zbraní, jak krátkými tak dlouhými. Tyto zbraně jsou rozděleny do 4 databází. První databáze byla tvořena v jiný den než databáze č. 2. Proto jsou odděleny. Důvodem rozdělení jsou podmínky při tvorbě. Jelikož výstřely byly zaznamenány mikrofonními jednotkami, mohlo dojít k různým nežádoucím vlivům. Například střelby v druhém střelišti nebo vliv klimatizace. Zbylé dvě databáze jsou tvořeny z dvou předchozích databází. Jsou děleny podle druhu a typu zbraně. Tyto dvě databáze jsou tvořeny proto, aby bylo zjištěno, jaký vliv na umělou neuronovou síť mají vzorky, které jsou rozděleny ve složkách podle názvu nebo když jsou různé zbraně v jedné složce.

Všechny databáze jsou uvedeny v tabulkách, kde je vidět počet vzorů popř. typ, druh zbraně, formát a frekvence.

Vzorky v databázi č. 1 jsou snímány ze dvou mikrofonních jednotek rozmístěných ve střelišti. Takže jsou dva druhy snímaných dat a to z kanálu 1 (CH00), který je ve vzdálenosti cca 17m a z kanálu 2 (CH01), který je ve vzdálenosti cca 20m. Proto jsou vzorky rozděleny podle kanálu nikoli rozděleny na polovinu. Úspěšnost klasifikace bude zkoušena s daty z jednoho kanálu nebo z druhého a to proto jestli je vzdálenost mikrofonních jednotek určující při vyhodnocování. U dalších databází už jsou vzorky jen z jedné mikrofonní jednotky.

První databáze obsahuje 5 zbraní. Z toho jsou 4 krátké a 1 dlouhá. První pokus byl závislý právě na této databázi, zde bude testováno, propojení této databáze a nastavení parametrů sítě. Zde je první tabulka, která představuje první učící a trénovací množinu dat.

#### **Pokus č. 1 s trénovací množinou dat z databáze č. 1(A, B)**

<i>Zbraň</i>	<i>Počet vzorů</i> <i>A (B)</i>	<i>Typ</i> <i>zbraně</i>	<i>Druh</i> <i>zbraně</i>	<i>Formát</i>	<i>Vzorkovací</i> <i>kmitočet</i>
<b>Arminius HW 38</b>	64(32)	krátká	revolver	wav	20 000 Hz
<b>CZ75SP 01</b>	68(34)	krátká	pistol	wav	20 000 Hz
<b>STI Spartan</b>	69(35)	krátká	pistol	wav	20 000 Hz
<b>Tokarev TT 33</b>	204(102)	krátká	pistol	wav	20 000 Hz
<b>Brokovnice</b>	5(2)	dlouhá	pistol	wav	20 000 Hz

*Tabulka 1: Přehled vzorů v databázi 1*

V následující tabulce jsou vzorky z databáze zvuků č. 2. Zde je zbraní více než v databázi č. 1. Uvidím jaký vliv na učení sítě má výstupní vrstva. Zde výstupní vrstva bude obsahovat 16 klasifikačních tříd. I zde snížím počet vzorů na polovinu v dalším testovacím cyklu.

**Pokus č. 2 s trénovací množinou dat z databáze č. 2(A, B)**

<i>Zbraň</i>	<i>Počet vzorů A(B)</i>	<i>Typ zbraně</i>	<i>Druh zbraně</i>	<i>Formát</i>	<i>Vzorkovací kmitočet</i>
<b>Colt 1873</b>	8(4)	krátká	revolver	wav	20 000 Hz
<b>Colt SAA</b>	5(2)	krátká	revolver	wav	20 000 Hz
<b>K100</b>	38(19)	krátká	pistol	wav	20 000 Hz
<b>Pistole 38</b>	20(10)	krátká	pistol	wav	20 000 Hz
<b>Saiga 308</b>	19(9)	dlouhá	pistol	wav	20 000 Hz
<b>SW27</b>	40(20)	krátká	revolver	wav	20 000 Hz
<b>SW629</b>	67(34)	krátká	revolver	wav	20 000 Hz
<b>Škorpion</b>	27(13)	dlouhá	samopal	wav	20 000 Hz
<b>Thompson</b>	33(17)	dlouhá	samopal	wav	20 000 Hz
<b>Tikka T3</b>	20(10)	dlouhá	pistol	wav	20 000 Hz
<b>SA58</b>	25(12)	dlouhá	samopal	wav	20 000 Hz
<b>Winchester 1866</b>	65(33)	dlouhá	pistol	wav	20 000 Hz

*Tabulka 2: Přehled vzorů v databázi 2*

Tabulka č. 3 dělí zbraně na krátké a dlouhé. Vzorky do této databáze jsem vybrala jak z databáze 1, tak z databáze 2.

### Pokus č. 3 s trénovací množinou dat z databáze č. 3

<i>Typ zbraně</i>	<i>Počet vzorů (A,B)</i>	<i>Formát</i>	<i>Vzorkovací kmitočet</i>
<b>Krátká</b>	583(291)	wav	20 000 Hz
<b>Dlouhá</b>	194(97)	wav	20 000 Hz

*Tabulka 3: Přehled vzorů v databázi 3*

Tabulka č. 4 obsahuje rozdělení vzorů podle typu zbraně. Do složky pistol jsem odebrala několik vzorů z první databáze a několik z databáze druhé. V tomto pokusu bude vidět, jak je síť schopna rozlišit vzorky i když nejsou rozděleny podle názvu, ale podle typu.

### Pokus č. 4 s trénovací množinou dat z databáze č. 4

<i>Typ zbraně</i>	<i>Počet vzorů (A,B)</i>	<i>Formát</i>	<i>Vzorkovací kmitočet</i>
<b>Pistol</b>	508(254)	wav	20 000 Hz
<b>Revolver</b>	184(92)	wav	20 000 Hz
<b>Samopal</b>	85(42)	wav	20 000 Hz

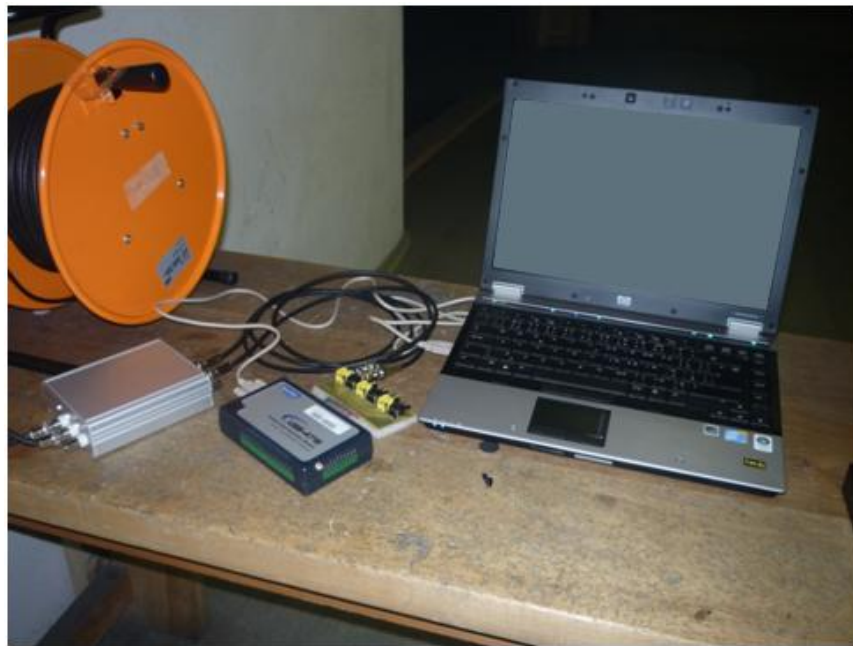
*Tabulka 4: Přehled vzorů v databázi 4*

### Použité zbraně

Na vytvoření databáze jsme využili zbraně, které jsou v tabulkách. U každé zbraně uvádím jen parametry, které jsou důležité pro neuronovou síť. Zbraně jsou vybírány podle možností střelnice.

## 4.2 Použité měřicí přístroje

Na to aby byla vytvořena kvalitní testovací a učící databáze výstřelů je potřeba také kvalitní techniky na měření a na vyhodnocení. Jako první je potřeba, aby byl analogový zvuk řádně zaznamenán mikrofonom. K dispozici jsou 2 mikrofony (viz. Kapitola 4.2.1 Mikrofonní jednotky). Tyto mikrofony byly rozmístěny na stojanech ve střelnici. K těmto mikrofonním jednotkám byla připojena jednotka pro sběr dat, která nám výstřely digitalizovala (viz. Kapitola 4.2.2 Jednotka pro sběr dat Advantech USB-4716).



Obrázek 14: Propojení snímací techniky a PC

### 4.2.1 Mikrofonní jednotky

Zvukové pole bylo snímáno mikrofonními jednotkami od firmy Brüel & Kjær, patřící mezi přední výrobce snímačů a vyhodnocovacích přístrojů pro měření hluku a vibrací. Pro aplikaci lokalizačního systému byly vybrány všesměrové mikrofonní jednotky typu 4188 doplněné předzesilovači typu 2671. Základní vlastnosti mikrofonních jednotek a předzesilovačů jsou uvedeny v tabulkách : *Tabulka 5: Parametry mikrofону B&K Type 4188 a Tabulka 6: Parametry předzesilovače B&K Type 2671.*



Obrázek 15: Mikrofon B&amp;K Type 4188

Průměr mikrofonu	1/2"
Dynamický rozsah	15.8 - 146 dB
Frekvenční rozsah	8 - 12500 Hz
Směrová charakteristika	všesměrová
Citlivost	31.6 mV/Pa
Polarizační napětí	0V (předpolarizován)

Tabulka 5: Parametry mikrofonu B&amp;K Type 4188



Obrázek 16: Předzesilovač B&amp;K Type 2671

Frekvenční rozsah	20 - 50000 Hz (+0,2dB; -2dB)
Vstupní impedance	1,5 GΩ
Výstupní impedance	< 50 Ω
Zkreslení	< -70 dB při 1.0 V <sub>out</sub> , 1 kHz
Napájení	DeltaTron 2 až 20mA (typ. 4mA)
Maximální DC výstupní napětí	12 V ±2 V
Typ konektoru	BNC

Tabulka 6: Parametry předzesilovače B&amp;K Type 2671

Napájení předzesilovačů konstantním proudem zajišťuje 3 kanálový DeltaTron zdroj ZU-T3T1 vybavený integrovaným akumulátorem umožňující jeho efektivní použití při měření v terénu.



*Obrázek 17: Zdroj proudu*

*DeltaTron – 3 kanály*

#### **4.2.2 Jednotka pro sběr dat Advantech USB-4716**

Pro digitalizaci vstupního analogového signálu z mikrofonních jednotek je použita přenosná jednotka pro sběr dat Advantech USB-4716. Její robustní konstrukce při udržení kompaktních rozměrů ji přímo předurčuje pro aplikace mobilního sběru dat. Díky podpoře technologie „Plug and Play“ je po připojení do USB portu počítače ihned automaticky rozpoznána a připravena k činnosti (za předpokladu, že je již v počítači nainstalován její ovladač a obslužné programové vybavení pro správu zařízení).

Jednotka je vybavena 16 analogovými vstupy (8 vstupy v případě diferenciálního režimu) s programově nastavitelnou vstupní citlivostí v několika pevně nastavených krocích v rozmezí od  $\pm 0,625$  V do  $\pm 10$  V. Vstupní impedance dosahuje hodnoty 1 G $\Omega$ . Vstupy mohou být nakonfigurovány do unipolárního nebo bipolárního režimu dle požadavků konkrétní aplikace. Rozlišení A/D převodníku je 16 bitů, maximální dosažitelná vzorkovací frekvence je 200 kHz. V případě vícekanálového režimu je tedy nutné tuto maximální hodnotu dělit počtem aktuálně používaných kanálů. Kromě analogových vstupů je k dispozici také 8 digitálních vstupů pracujících s logickými úrovněmi 3,3V / TTL.

Na výstupní straně je jednotka vybavena dvěma analogovými výstupy s rozlišením 16 bitů, které lze přepnout na napěťové rozsahy 0 ~ 5 V, 0 ~ 10 V,  $\pm 5$  V a  $\pm 10$  V. Dále je implementováno 8 digitálních TTL výstupů s proudovou zatížitelností až 6 mA..

Kód zesílení		4	0	1	2	3
Zesílení		0,5	1	2	4	8
Rozsah	Bipolární	$\pm 10$ V	$\pm 5$ V	$\pm 2,5$ V	$\pm 1,25$ V	$\pm 0,625$ V
	Unipolární	---	0 ~ 10 V	0 ~ 5 V	0 ~ 2,5 V	0 ~ 1,25 V

Tabulka 7: Vstupní rozsahy jednotky USB-4716



Obrázek 18: Mobilní jednotka pro sběr dat

*Advantech USB-4716*

Tato snímací a zaznamenávací technika byla použita při měření v terénu. Jako zdroj informací pro tento popis byly přiložené manuály.

### 4.3 Popis uživatelské aplikace

Jedním z výstupů diplomové práce je uživatelský program, který umožňuje klasifikovat neznámé výstřely do zadaných klasifikačních tříd použitím neuronové sítě.

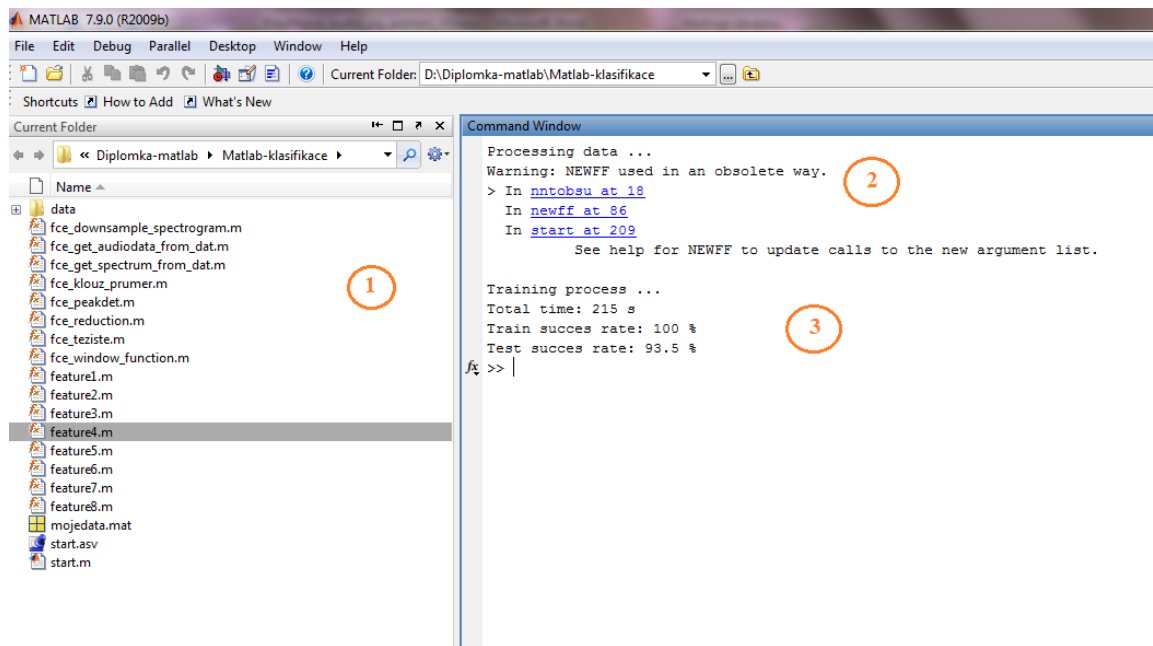
Uvedená úloha je řešena s využitím programového produktu MATLAB 7.9.0 (R2009b).

Na to aby program vznikl, potřebujeme uskutečnit tyto kroky:

- Sběr dat (na trénování a učení sítě)
  - Vytvoření databáze, změna formátu, normalizace
- Vytvoření sítě (algoritmus)
  - Vlastní program v programovém prostředí Matlab
- Nakonfigurování sítě (určit počet vrstev, počet neuronů)
  - Volba topologie je hlavní etapa celého procesu klasifikace. Navrhne se struktura vlastní sítě, počet skrytých vrstev a počet neuronů v těchto vrstvách.
- Určení vah
  - Váhy si program určí sám (náhodné nastavení)
- Trénování sítě
  - Na trénování je použit backpropagation algoritmus tzv. Levenbergův-Marquardtův algoritmus.
- Využití sítě

#### 4.3.1 Grafické zobrazení aplikace

Grafický náhled do programu Matlab. Tato hlavní část se skládá z dvou pracovních oken. První okno Current folder zaznamenává zdrojovou cestu programu. Nejen programu ale také dat, které slouží jako úložiště trénovací množiny. Druhé okno už poukazuje na proces, který se vykonává spuštěním programu. Popisek číslo 2 ukazuje, že program zpracovává data, tzn., že na vzory aplikuje daný feature vector jinak řečeno extrahuje charakteristické rysy ze signálu. Tyto extrahované signály jsou redukovány a nachystány pro vstup do neuronové sítě jako trénovací, validační nebo testovací množina dat. Poslední popisem v příkazovém okně (Command Windows) je trénovací proces, kde už samotná neuronová síť trénuje. S tímto procesem je spojen následující obrázek Grafická podoba trénovacího procesu, který ukazuje, v jakém stádiu se trénování umělé neuronové sítě nachází.



Obrázek 19: Program Matlab

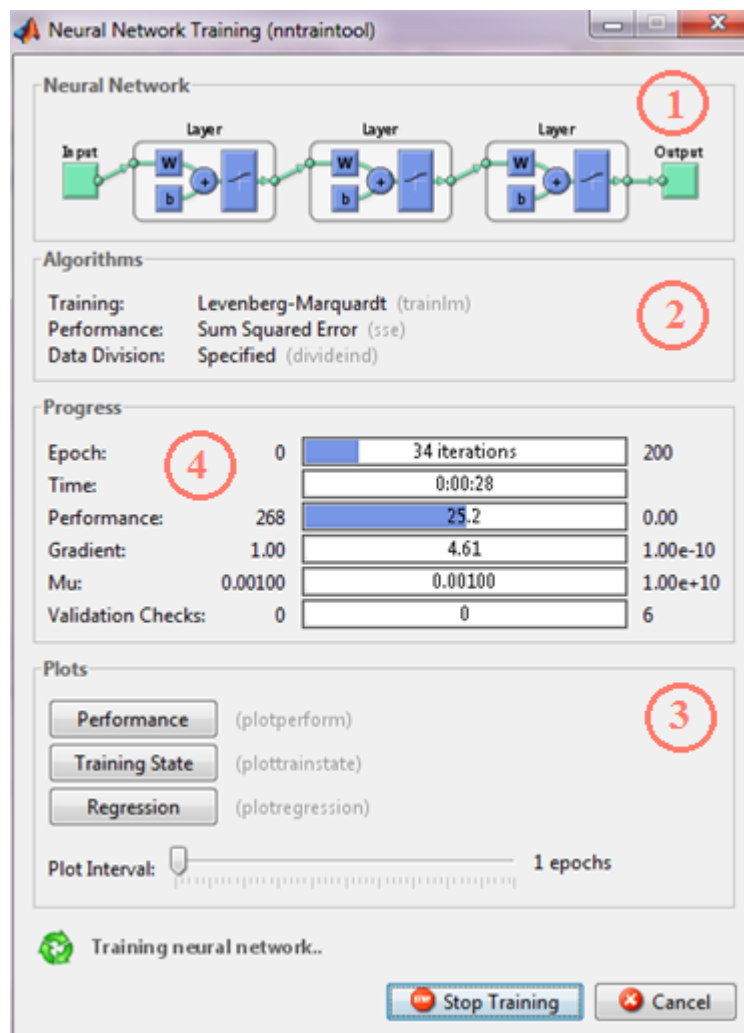
### 4.3.2 Proces trénování

Následující obrázek č. 20 ukazuje průběh trénování sítě. Je to vlastně grafické okno průběhu trénování umělé neuronové sítě. Z tohoto okna jsme schopni vyčíst kolik je nastaveno skrytých vrstev toto lze vidět v popisu č.1. V případě tohoto trénování jsou nastaveny 3 skryté vrstvy. Pod číslem 2 se skrývá informace o algoritmu, který je použit pro neuronovou síť. Jako trénovací funkce je v algoritmu nastaven Levenberg – Marquardt. Performance udává chybu trénování, zde je nastaven součet chyb (sse – Sum Squared Error).

Okno 4 udává, za jak dlouho se síť natrénovala a s kolika iteracemi (epochami). Performance udává chybu trénování. Posledním důležitým bodem v tomto okně je Validation Checks což je kritérium pro ukončení trénování sítě. [29]

Vývoj – postup učení

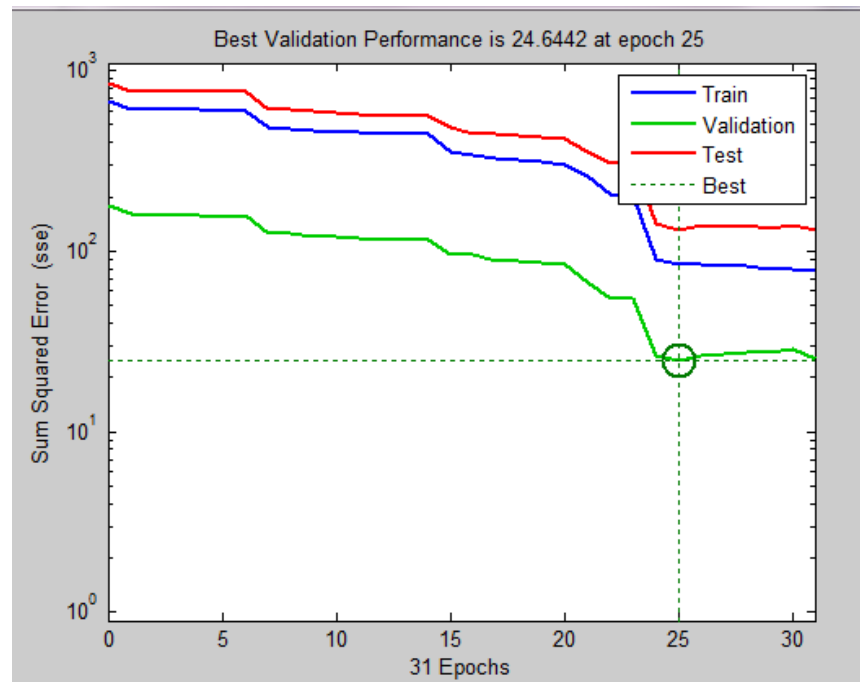
- Epoch – počet učících kroků ( předkládání vzorů k trénování)
- Time - maximální doba, kdy bude síť učena
- Performance – udává chybu trénování
- Gradient – směrnice k výsledku
- Mu – střední chyba
- Validation checks – zajišťuje, aby se neuronová síť nepřeučila



Obrázek 20: Grafická podoba trénovacího procesu

Proces učení probíhá tak, že se vstupní signály opakovaně v jednotlivých epochách přikládají na vstup neuronové sítě a výstupní vzory se porovnávají s výstupem z neuronové sítě. Cílem učení je pak postupnou změnou vah spojení mezi jednotlivými neurony, případně strmosti aktivačních funkcí, docílit po určitém počtu epoch trénování, minimálního rozdílu mezi výstupem ze soustavy) a výstupem z neuronové sítě [13]

Poslední okno č. 3 v grafické podobě trénovacího procesu, jsou grafy, které zobrazují průběh trénování sítě (plots)



Obrázek 21: Závislost trénovací chyby na počtu epoch

Graf poskytuje nahlédnutí, jak se mění závislost součtu chyb (SSE – Sum Squared Error) na počtu epoch, spolu s doplňujícími informacemi o Gradientu nebo Mu, či Validation Checks.[15] Pokud validace (zelená) dosáhne středu, tak je síť naučena a dojde k ukončení trénování.

#### 4.4 Algoritmus

Naprogramování samotné umělé neuronové sítě je velmi obtížné, proto jsem využila stávající program a obohatila tento program o další funkce. A to definování charakteristik neboli feature vector.

Charakteristiky jsou celkem 4, a jsou popsány výše v kapitole 3.2. U těchto charakteristik se nastavují parametry pro každou charakteristiku jiné.

K realizaci algoritmu neuronové sítě existuje několik příkazů. Pokud chceme dopřednou neuronovou síť, použijeme příkaz `newff()`. Před dalším krokem, což je trénování sítě, je zapotřebí inicializace. To znamená nastavení vah a prahů neuronové sítě `net = init(net)`; Potom už následuje vlastní proces trénování. Ten se spouští příkazem `train()`, který má několik argumentů, definujících trénovací, validační a testovací množinu. K simulaci sítě slouží příkaz `sim()`.

Vzorky jsou rozděleny 60% na trénování, validace 10% a na testování 40%. Samozřejmě i tyto parametry jsou v programu měnitelné. Já ponechala rozdělení v tomto tvaru. 60% z vzorů na trénování, podle těchto vzorů se UNS trénuje. Validace slouží k zabránění přeučení sítě a na to stačí právě 10%. 40% na samotné testování sítě.

#### 4.5 Úvod do experimentální práce

Experimentální činnost v této diplomové práci představovala hledání optimálního nastavení parametrů sítě (nastavení počtu neuronů ve skrytých vrstvách). Dále jaká extrakce charakteristického rysu ať už pomocí průměrovaného spektra, autokorelační funkce, akustického tlaku nebo centrálních momentů je nejvhodnější na zredukování signálu. Právě tyto analýzy jak frekvenční, tak časové v této práci hrají důležitou úlohu, jelikož má každá jiné parametry.

Při návrhu topologie celé sítě je jedním ze stěžejních úkonů volba počtu skrytých vrstev. Universální řešení neexistuje, každá praktická situace si žádá nastavení vrstev a neuronů k dosažení uspokojivých výsledků. Nicméně, zkušenosti ukazují [13], že vyhovujících výsledků se dá dosáhnout v mnoha případech pomocí dvouvrstvé neuronové sítě. Já jsem si v této práci stanovila počet neuronů a skrytých vrstev pevně a to v první skryté vrstvě 5, 10, 12, 15. Ve vrstvě druhé 2 neurony. S těmito hodnotami budu pracovat (měnit a kombinovat) a zjišťovat jak se mění výsledná klasifikace. Tento počet je zvolen na základě možností mého osobního počítače.

Doba trénování také závisela na počtu vstupních dat, proto jsou databáze rozděleny na dvě, A a B. V databázi B je počet vzorů poloviční než v databázi A. Tímto bude zjištěno jaký vliv na klasifikaci má počet vstupních vzorů.

Akorát vzorky v databázi č. 1 jsou snímány ze dvou mikrofonních jednotek rozmístěných ve střelišti. Takže jsou dva druhy snímaných dat a to z kanálu 1 (CH00), který je ve vzdálenosti cca 17m a z kanálu 2 (CH01), který je ve vzdálenosti cca 20m. Proto jsou vzorky rozděleny podle kanálu nikoli rozděleny na polovinu. Úspěšnost klasifikace bude zkoušena s daty z jednoho kanálu nebo z druhého a to proto jestli je vzdálenost mikrofonních jednotek určující při vyhodnocování. U dalších databází už jsou vzaty vzorky jen z jedné mikrofonní jednotky.

Každá charakteristika má své parametry, které je možné v algoritmu měnit a zkoušet jaká je úspěšnost. Jednotlivé parametry jsou popsány u každé charakteristiky v kapitole **Chyba! Nenalezen zdroj odkazů..**

Všechny pokusy klasifikace jsou prováděny 6x. V tabulkách jsou uváděny nejlepší úspěšnosti při testování/trénování.

#### 4.5.1 Popis tabulky

Tato tabulka je výsledkem klasifikačního cyklu pomocí neuronových sítí. První sloupec Databáze uvádí, která databáze je na vstupu. Jsou celkem 4 databáze, které jsou rozděleny na A i B. Obsah databází je detailně popsán v kapitole **Chyba! Nenalezen zdroj odkazů.** Sloupec s názvem Vstupní vrstva obsahuje počet vstupů do neuronové sítě, tyto vstupy jsou dány nastavením parametrů u dané charakteristiky. Tato analýza, zredukuje vstupní signál podle svých parametrů viz kapitola 3.2 Výběr charakteristik. Sloupec Výstupní vrstva je dána počtem klasifikačních tříd. Tyto třídy jsou v podstatě složky v databázi. V databázi č. 1 je 5 zbraní, každá ve své složce, proto výstupem je 5 neuronů. Jak v části Topologie neuronových sítí je uvedeno tato síť je dvouvrstvá, tzn., že má 2 skryté vrstvy. Sloupec, který uvádí hodnoty [5;2] se nazývá Počet neuronů. Tento sloupec poukazuje na skryté vrstvy, tzn., že v první skryté vrstvě je 5 neuronů a 2 neurony jsou v druhé skryté vrstvě. Další sloupec je proměnlivý, záleží na tom, jaká charakteristika je použita. U této tabulky, která je vzata z testovacího cyklu č. 1 u průměrovaného spektra, je dalším sloupcem Průměrovací krok. Jiné analýzy mají například redukční faktor, počet momentů apod. Tyto parametry se nastavují přímo v jádru programu u každé analýzy. Poslední sloupec uvádí procentuální úspěšnost neuronové sítě. Jak byla schopna klasifikovat vzorky do klasifikačních tříd. 96/77% uvádí procentuální úspěšnost jak trénování/testování.

Databáze	Vstupní vrstva	Výstupní vrstva	Počet neuronů	Průměrovací krok	Úspěšnost
1A	238	5	[5;2]	10	<b>96/77%</b>
1A	238	5	[12;2]	10	<b>100/89%</b>
1B	238	5	[10;2]	10	<b>86/80%</b>
1B	238	5	[15;2]	10	<b>100/96%</b>
1A	119	5	[5;2]	20	<b>99/90%</b>
1A	119	5	[12;2]	20	<b>100/95%</b>
1B	119	5	[10;2]	20	<b>82/79%</b>
1B	119	5	[15;2]	20	<b>49/40%</b>

Tabulka 8: Tabulka s legendou

**Obsah databáze zvuků 1A, 1B**

Tato tabulka uvádí počet vzorů v databázi č. 1. Databáze je rozdělena podle vzdálenosti snímání zvuku. První mikrofonní jednotka byla ve vzdálenosti 17 m CH00, druhá mikrofonní jednotka byla ve vzdálenosti 20 m CH01. Přesný obsah databáze( název zbraně, typ zbraně) v kapitole 4.1.1.

<i>Zbraň</i>	<i>Počet vzorů</i> <i>A (B)</i>	<i>Typ</i> <i>zbraně</i>	<i>Druh</i> <i>zbraně</i>
<b>Arminius HW 38</b>	64(32)	krátká	revolver
<b>CZ75SP 01</b>	68(34)	krátká	pistol
<b>STI Spartan</b>	69(35)	krátká	pistol
<b>Tokarev TT 33</b>	204(102)	krátká	pistol
<b>Brokovnice</b>	5(2)	dlouhá	pistol

*Tabulka 9: Popis databáze výstřelů*

Celkem v databázi 1A je 410 výstřelů, v databázi vzorů 1B je 205 výstřelů z výše jmenovaných zbraní.

Jak je vidět průměrované spektrum zredukovalo počet vzorů cca na polovinu vzorů z originální databáze zvuků. Tato redukce je dána u této analýzy průměrovacím krokem.

## 4.5.2 Testovací cyklus č.1

Průměrované spektrum

Pokusy byly měřeny při těchto podmínkách:

*Minimální frekvence: 100 Hz*

*Maximální frekvence: 3000 Hz*

*FFT: 16 384<sup>4</sup>*

*Průměrovací krok: 10, 20*

## Testovací a trénovací cyklus

Databáze	Vstupní vrstva	Výstupní vrstva	Počet neuronů	Průměrovací krok	Úspěšnost
1A	238	5	[5;2]	10	<b>96/77%</b>
1A	238	5	[12;2]	10	<b>100/89%</b>
1B	238	5	[10;2]	10	<b>86/80%</b>
1B	238	5	[15;2]	10	<b>100/96%</b>
1A	119	5	[5;2]	20	<b>99/90%</b>
1A	119	5	[12;2]	20	<b>100/95%</b>
1B	119	5	[10;2]	20	<b>82/79%</b>
1B	119	5	[15;2]	20	<b>49/40%</b>

*Tabulka 10: Učící a trénovací cykly s použitím průměrového spektra*

Úspěšnost trénování sítě je zaznamenána taková, která byla nejlepší po šesti pokusech. Databáze 1A obsahuje vzorky z CH00. Databáze 1B obsahuje vzorky z CH01 a při procentuální úspěšnosti, jak lze z tabulky vidět se na databázi 1B síť učila o pár procent lépe než na databázi, kde jsou vzorky z CH00. Průměrovací krok je nastaven na 10 a 20, podle tabulky je zřejmé, že průměrovací krok 10 vykazuje lepší klasifikační shodu.

<sup>4</sup> Počet bodů ze signálu brané k frekvenční analýze

**Testovací a trénovací cyklus**

Databáze	Vstupní vrstva	Výstupní vrstva	Počet neuronů	Průměrovací krok	Úspěšnost
2A	238	12	[5;2]	10	<b>70/53%</b>
2A	238	12	[12;2]	10	<b>66/62%</b>
2B	238	12	[10;2]	10	<b>64/62%</b>
2B	238	12	[15;2]	10	<b>57/40%</b>
2A	119	12	[5;2]	20	<b>59/47%</b>
2A	119	12	[12;2]	20	<b>28/25%</b>
2B	119	12	[10;2]	20	<b>9/5%</b>
2B	119	12	[15;2]	20	<b>49/44%</b>

*Tabulka 11: Učící a trénovací cykly s použitím průměrového spektra*

Při nastavení, kde je 238 vstupních vzorů vykazovala síť velmi rozdílné a nízké úspěšnosti. Databáze č. 2 pro tuto charakteristiku s tímto nastavením neuronové sítě nevykazuje velké úspěšnosti. Možností proč úspěšnost je, tak malá je mnoho například to může být způsobeno mnoha vzory v databázi nebo také, že nastavení neuronů je nevhodné pro tolik vzorů.

Při zvyšování průměrovaného kroku, úspěšnost klasifikace klesala. Ale při velkém snižování klesala také. Optimální nastavení průměrovaného kroku záviselo taky na nastavení počtu neuronů.

**Testovací a trénovací cyklus**

Databáze	Vstupní vrstva	Výstupní vrstva	Počet neuronů	Průměrovací krok	Úspěšnost
3A	238	2	[5;2]	10	<b>100/99%</b>
3A	119	2	[12;2]	20	<b>100/99%</b>
3B	238	2	[5;2]	10	<b>100/95%</b>
3B	119	2	[12;2]	20	<b>100/97%</b>

*Tabulka 12: Učící a trénovací cykly s použitím průměrového spektra*

Jako nejlepší nastavení sítě na databázi 3B je průměrovací krok 20 a počet neuronů ve skrytých vrstvách 12,2. U této topologie všechny 4 pokusy byly nad 95%. Zase u databáze 3A úspěšnost vycházela 100/99% u obou nastaveních.

Celkově tato databáze vykazovala velmi dobré výsledky jak v učení, tak v trénování sítě. Důvod této velké úspěšnosti může být v klasifikačních třídách, jelikož v této databázi jsou jen 2 klasifikační třídy. Oproti předchozím databázím je zde rozdíl v tom, že různé zbraně jsou v jedné složce (krátká, dlouhá). Jsou zde akorát 2 výstupní klasifikační třídy.

**Testovací a trénovací cyklus**

Databáze	Vstupní vrstva	Výstupní vrstva	Počet neuronů	Průměrovací krok	Úspěšnost
4A	238	3	[5;2]	10	<b>99/94%</b>
4A	119	3	[12;2]	20	<b>100/97%</b>
4B	238	3	[5;2]	10	<b>100/93%</b>
4B	119	3	[12;2]	20	<b>100/94%</b>

*Tabulka 13: Učící a trénovací cykly s použitím průměrového spektra*

U databáze 4, kde jsou vzorky rozděleny podle typu zbraně je vidět, že úspěšnost u této frekvenční analýzy je vysoká. Tato úspěšnost by mohla být přisuzována výstupními třídami, kde jak v databázi č. 3, tak v databázi č 4 jsou pouze 2 nebo 3 klasifikační třídy a vzorky nejsou rozděleny do mnoha složek.

## 4.5.3 Testovací cyklus č. 2

Autokorelační funkce

*Pokusy byly měřeny při těchto podmínkách:*

*Minimální frekvence: 100 Hz*

*Maximální frekvence: 5000 Hz*

*FFT: 16 384*

*Redukční faktor: 10, 20, 40*

*Klouzavý průměr: 7, 9, 11, 17*

## Testovací a trénovací cyklus

Databáze	Vstupní vrstva	Výstupní vrstva	Počet neuronů	Klouzavý průměr	Redukční faktor	Úspěšnost
1A	402	5	[5;2]	7	10	<b>81/79%</b>
1A	201	5	[12;2]	7	20	<b>82/82%</b>
1B	101	5	[10;2]	9	40	<b>84/78%</b>
1B	101	5	[15;2]	9	40	<b>97/95%</b>
1A	402	5	[5;2]	11	10	<b>76/73%</b>
1A	201	5	[12;2]	11	20	<b>82/82%</b>
1B	101	5	[10;2]	17	40	<b>100/96%</b>
1B	101	5	[15;2]	17	40	<b>100/97%</b>

*Tabulka 14: Učící a trénovací cykly s použitím autokorelační funkce*

Jak je vidět v tabulce, tak autokorelační funkce vykazuje nejlepší výsledky s databází vzorů, 1B a s větším klouzavým průměrem, tak redukčním faktorem 40. Vyšší redukční faktor už vykazoval pokles úspěšnosti v trénování/testování sítě.

**Testovací a trénovací cyklus**

Databáze	Vstupní vrstva	Výstupní vrstva	Počet neuronů	Klouzavý průměr	Redukční faktor	Úspěšnost
2A	402	12	[5;2]	7	10	<b>42/39%</b>
2A	201	12	[12;2]	7	20	<b>68/49%</b>
2B	101	12	[10;2]	9	40	<b>61/63%</b>
2B	101	12	[15;2]	11	40	<b>57/58%</b>
2A	402	12	[5;2]	11	10	<b>35/32%</b>
2A	201	12	[12;2]	11	20	<b>77/56%</b>
2B	101	12	[10;2]	17	40	<b>58/39%</b>
2B	101	12	[15;2]	17	40	<b>67/52%</b>

*Tabulka 15: Učící a trénovací cykly s použitím autokorelační fce*

Sít s touto analýzou a databází vykazovala velmi rozdílné výsledky. Nejméně vhodné nastavení počtu neuronů je [5;2] u obou databází. Nejlepší by se dalo vyhodnotit [12;2] nastavení. Jak u databáze 2A tak 2B. Při snižování redukčního faktoru rapidně stoupal čas, za který se UNS trénovala a úspěšnost byla okolo 20%.

**Testovací a trénovací cyklus**

Databáze	Vstupní vrstva	Výstupní vrstva	Počet neuronů	Klouzavý průměr	Redukční faktor	Úspěšnost
3A	202	2	[5;2]	9	20	<b>99/93%</b>
3A	101	2	[12;2]	17	40	<b>99/98%</b>
3B	202	2	[5;2]	9	20	<b>78/70%</b>
3B	101	2	[12;2]	17	40	<b>98/94%</b>

*Tabulka 16: Učící a trénovací cykly s použitím autokorelační fce*

Jak v předchozí analýze průměrované spektrum databáze 3 vykazuje velmi dobré výsledky.

**Testovací a trénovací cyklus**

Databáze	Vstupní vrstva	Výstupní vrstva	Počet neuronů	Klouzavý průměr	Redukční faktor	Úspěšnost
4A	202	3	[5;2]	9	20	<b>64/68%</b>
4A	101	3	[12;2]	17	40	<b>86/83%</b>
4B	202	3	[5;2]	9	20	<b>49/46%</b>
4B	101	3	[12;2]	17	40	<b>49/47%</b>

*Tabulka 17: Učící a trénovací cykly s použitím autokorelační funkce*

Více neuronů, větší klouzavý průměr a velký redukční faktor jak se zdá je pro tento cyklus nejvíce vyhovující.

## 4.5.4 Testovací cyklus č. 3

Akustický tlak

*Pokusy byly měřeny při těchto podmínkách:*

*Redukční faktor: 20, 30, 40, 50*

*Délka signálu: 1.0, 1.2, 1.4 s*

*Posun: 0,1 s*

**Testovací a trénovací cyklus**

Databáze	Vstupní vrstva	Výstupní vrstva	Počet neuronů	Délka signálu (s)	Redukční faktor	Úspěšnost
1A	500	5	[5;2]	1.0	40	<b>59/60%</b>
1A	800	5	[5;2]	1.2	30	<b>45/44%</b>
1B	500	5	[5;2]	1.0	40	<b>60/66%</b>
1B	480	5	[5;2]	1.4	50	<b>78/80%</b>
1A	500	5	[5;2]	1.4	40	<b>55/54%</b>
1A	800	5	[5;2]	1.0	30	<b>44/44%</b>
1B	500	5	[5;2]	1.2	40	<b>61/62%</b>
1B	480	5	[5;2]	1.0	50	<b>80/78%</b>

*Tabulka 18: Učící a trénovací cykly s použitím akustického tlaku*

Při pokusech s redukováným akustickým tlakem byl problém, jelikož vyšší nastavení počtu neuronů jak [5;2] ve skrytých vrstvách docházelo k přerušení trénování kvůli času, za který se síť měla naučit. Tento čas někdy byl dokonce až 2 hodiny. Proto je topologie volena pouze jedna a to právě 5 neuronů v první skryté vrstvě a 2 neurony v druhé skryté vrstvě. Pak už se čas pohyboval okolo 30 minut. Jak je vidět zvýšila jsem i redukční faktor, i přesto to bylo velmi časově náročné oproti jiným pokusům, kdy se čas pohyboval okolo 1min.

Úspěšnost tohoto testovacího cyklu nebyla nějak závratná, ale v průměru na 50% vyšla vždy.

**Testovací a trénovací cyklus**

Databáze	Vstupní vrstva	Výstupní vrstva	Počet neuronů	Délka signálu (s)	Redukční faktor	Úspěšnost
2A	500	12	[5;2]	1.0	40	<b>89/88%</b>
2A	800	12	[5;2]	1.2	30	<b>82/82%</b>
2B	500	12	[5;2]	1.0	40	<b>98/86%</b>
2B	480	12	[5;2]	1.4	50	<b>99/90%</b>
2A	500	12	[5;2]	1.4	40	<b>84/80%</b>
2A	800	12	[5;2]	1.0	30	<b>65/64%</b>
2B	500	12	[5;2]	1.2	40	<b>90/89%</b>
2B	480	12	[5;2]	1.0	50	<b>80/85%</b>

*Tabulka 19: Učící a trénovací cykly s použitím akustického tlaku*

Tento cyklus vykazoval lepší úspěšnost než v předchozím cyklu s touto analýzou. Jednoznačná příčina je v kvalitě vzorků, jelikož nastavení počtu neuronů i parametrů je stejné jako v předchozím cyklu. Jak je vidět každá analýza má jiné nároky na vstupní vzorky. U průměrovaného spektra byla lepší úspěšnost s databází č.2 naopak autokorelační funkci tato databáze vyhovovala méně. Je zde také problém, že jsou topologie sítě jinak nastavené, takže nelze přesně porovnávat s předchozí analýzou.

**Testovací a trénovací cyklus**

Databáze	Vstupní vrstva	Výstupní vrstva	Počet neuronů	Délka signálu (s)	Redukční faktor	Úspěšnost
3A	1400	2	[5;2]	1.0	20	<b>75/76%</b>
3A	934	2	[5;2]	1.0	30	<b>100/96%</b>
3B	700	2	[5;2]	1.4	40	<b>100/97%</b>
3B	700	2	[5;2]	1.0	40	<b>100/97%</b>

*Tabulka 20: Učící a trénovací cykly s použitím akustického tlaku*

Databáze 3 i propojení neuronů jak je vidět v tabulce je jednou z nejlepších databází pro autokorelační funkci.

**Testovací a trénovací cyklus**

Databáze	Vstupní vrstva	Výstupní vrstva	Počet neuronů	Délka signálu (s)	Redukční faktor	Úspěšnost
4A	667	3	[5;2]	1.0	20	<b>66/70%</b>
4A	667	3	[5;2]	1.0	30	<b>88/83%</b>
4B	667	3	[5;2]	1.4	40	<b>99/85%</b>
4B	667	3	[5;2]	1.4	40	<b>87/90%</b>

*Tabulka 21: Učící a trénovací cykly s použitím akustického tlaku*

U akustického tlaku se jako nejlepší propojení jeví celkový čas signálu 1.4 s a redukční faktor 40 nebo 50. Ze 4 databází byla úspěšnost tohoto nastavení u 3 databází nad 90%.

## 4.5.5 Testovací cyklus č. 4

Centrální momenty

*Pokusy byly měřeny při těchto podmínkách:*

*Posun: 0.1 s*

*Délka signálu: 1.0, 1.2, 1.4 s*

*Počet momentů: 6, 8, 10, 12, 15,17*

## Testovací a trénovací cyklus

Databáze	Vstupní vrstva	Výstupní vrstva	Počet neuronů	Délka signálu (s)	Počet momentů	Úspěšnost
1A	6	5	[5;2]	1.2	6	<b>45/51%</b>
1A	8	5	[12;2]	1.4	8	<b>84/81%</b>
1B	6	5	[10;2]	1.0	6	<b>66/64%</b>
1B	6	5	[15;2]	1.0	6	<b>86/82%</b>
1A	8	5	[5;2]	1.4	8	<b>69/63%</b>
1A	6	5	[12;2]	1.2	6	<b>50/49%</b>
1B	8	5	[10;2]	1.0	8	<b>92/83%</b>
1B	8	5	[15;2]	1.4	8	<b>81/82%</b>

*Tabulka 22: Učící a trénovací cykly s použitím centrálních momentů*

Databáze 1B vykazuje lepší úspěšnost. U této databáze jde vidět, že počet momentů ovlivňuje výslednou úspěšnost.

**Testovací a trénovací cyklus**

Databáze	Vstupní vrstva	Výstupní vrstva	Počet neuronů	Délka signálu (s)	Počet momentů	Úspěšnost
2A	6	5	[5;2]	1.2	6	<b>40/35%</b>
2A	8	5	[12;2]	1.2	8	<b>45/37%</b>
2B	17	5	[10;2]	1.2	17	<b>17/18%</b>
2B	6	5	[15;2]	1.2	6	<b>19/18%</b>
2A	8	5	[5;2]	1.2	8	<b>50/46%</b>
2A	12	5	[12;2]	1.2	12	<b>22/20%</b>
2B	8	5	[10;2]	1.2	8	<b>21/20%</b>
2B	8	5	[15;2]	1.2	8	<b>28/27%</b>

*Tabulka 23: Učící a trénovací cykly s použitím centrálních momentů*

**Testovací a trénovací cyklus**

Databáze	Vstupní vrstva	Výstupní vrstva	Počet neuronů	Délka signálu (s)	Počet momentů	Úspěšnost
3A	20	2	[5;2]	1.2	10	<b>97/95%</b>
3A	8	2	[12;2]	1.2	8	<b>98/94%</b>
3B	30	2	[5;2]	1.2	15	<b>100/96%</b>
3B	8	2	[12;2]	1.2	8	<b>100/97%</b>

*Tabulka 24: Učící a trénovací cykly s použitím centrálních momentů*

**Testovací a trénovací cyklus**

Databáze	Vstupní vrstva	Výstupní vrstva	Počet neuronů	Délka signálu (s)	Počet momentů	Úspěšnost
4A	10	2	[5;2]	1.2	10	<b>89/88%</b>
4A	8	2	[12;2]	1.2	8	<b>89/87%</b>
4B	30	2	[5;2]	1.2	30	<b>80/78%</b>
4B	8	2	[12;2]	1.2	8	<b>82/80%</b>

*Tabulka 25: Učící a trénovací cykly s použitím centrálních momentů*

**Zhodnocení experimentální činnosti**

Na extrakci charakteristických rysů jsem použila 4 analýzy, jak z frekvenční oblasti, tak z časové. Z časového hlediska se nejlépe testovala síť s použitím analýzy centrální momenty, tato analýza vykazovala čas trénování do 5s. Avšak jak je vidět úspěšnost této charakteristiky a nastavením počtu neuronů nebyla, tak úspěšná jako například při trénování sítě za použití průměrovaného spektra. U této charakteristiky jsem dosáhla ve většině případů 100% při trénování. Časová náročnost u této charakteristiky byla do 1 minuty. U akustického tlaku doba trénování byla někdy, tak dlouhá, že docházelo k přerušení trénování. Z tohoto důvodu jsem se rozhodla nastavení počtu neuronů ve skrytých vrstvách snížit na [5;2], jelikož vyšší počet neuronů vykazoval zase přerušení učení.

Počet neuronů nelze stanovit, jelikož u každé charakteristiky ukazuje toto nastavení jinou úspěšnost.

Po experimentální činnosti jsem došla k závěru, že výsledky klasifikace jsou lepší, pokud je méně klasifikačních tříd. Tento fakt potvrzuje experimentální činnost s použitím databáze vzorů. 3 a č. 4, kde jsou jen 2 nebo 3 klasifikační třídy. Úspěšnost se pohybovala okolo 90%.

## ZÁVĚR

Tématem diplomové práce byla klasifikace audio signálu pomocí neuronových sítí. Toto téma je velmi obsáhlé, proto jsem se rozhodla popsat v teoretické části opravdu ty nejdůležitější části. Obeznámila jsem čitatele s problematikou klasifikace, historií a funkcí neuronových sítí a s problematikou zpracování audio signálů.

Stávající databáze, kterou jsem měla k dispozici, od svého vedoucího práce nebyla dostačující na tento projekt, proto jsme se rozhodli v rámci praktické části vytvořit novou databázi výstřelů. Tato tvorba probíhala v terénu a to přímo na střelnici v Otrokovicích. Střelba byla rozdělena do 2 dnů. Měli jsme k dispozici celé střeliště. Zbraní bylo celkem 17 druhů ať pistolí, revolverů či samopalů. Při tvorbě konkrétně střelbě ze zbraní byl kladen důraz na to, aby zvuky byly co nejlépe zaznamenány. Abychom dosáhli, co nejkvalitnějších zvuků byly k tomu použity kvalitní mikrofonní a snímací jednotky, které byly rozmístěny přímo ve střelišti.

V praktické části bylo prvním cílem vytvoření právě zmiňované databáze zvuků a dalším cílem bylo upravení stávajícího programu na klasifikaci signálů pomocí neuronové sítě.

Úprava algoritmu spočívala v doplnění funkcí na extrakci charakteristických rysů ze signálu. Na extrahování byly použity analýzy jak z frekvenční oblasti, tak z časové. Celkem z každé oblasti 2 charakteristiky. Více v kapitole 3.2. o těchto charakteristikách. Po úpravě algoritmu, která spočívala v doplnění podprogramů, na extrakci charakteristických rysů ze signálu, následovala experimentální činnost a to pokusit se najít takovou nastavení sítě jako je propojení vrstev a neuronů, tak aby klasifikace byla co nejuspěšnější. Nejlepší výsledky v mnoha případech potvrdily teorii v kapitole Experimentální práce, a sice to, že dvouvrstvá síť vykazuje dobré výsledky. Na úspěšnosti nezávisela jen správné nastavení sítě ale také právě analýze, která extrahovala charakteristické rysy ze signálu.

Jediný problém při zpracování této práce byl při experimentální činnosti. Konkrétně s časovou analýzou, kdy při nastavování počtu neuronů ve skrytých vrstvách u akustického tlaku. Nastavení neuronů u skrytých vrstev více jak [5;2] docházelo k přerušení trénování z důvodu dlouhého času trénování. Čas kolikrát přesahoval i 2 hodiny a úspěšnost klasifikace byla malá. Je možné, že je způsobeno výkonem mého osobního počítače nebo spíše i jak jde vidět v tabulce 16 vstupní vrstvou, která obsahuje mnoho vstupních vektorů. Doporučila bych snížit redukční faktor.

Výstupem je tedy upravený program v programovém prostředí MATLAB 2009 a vytvořená a upravená učící množina zvuků.

Úkolem bylo najít takové parametry neuronové sítě (počty neuronů, vrstev apod.) aby došlo, k co maximální shodě klasifikace což se v určitých případech podařilo.

Tato práce obsahuje experimentální činnost, která zjišťuje jaké nastavení sítě je nejvhodnější pro danou charakteristiku signálu s určitou databází vzorů. Jako návrh v pokračování bych doporučila určitě více druhů vzorů do databáze. Tyto vzorky bych se snažila tvořit v odlišných prostředích. Například na střelnici v otevřeném prostředí. Dále bych vyzkoušela další analýzy signálu například spektrogram, která já jsem použít v této práci nemohla, jelikož mi to výkonnost osobního počítače nedovolila.

Parametry mého počítače jsou MS Windows 7 Home Premium 64Bit. Procesor Pentium Dual-Core CPU T4300 2.10GHz, RAM o velikosti 3.0GB. Parametry uvádím proto, že na kombinace některých nastavení sítě byla tak náročná, že trvalo dokonce i hodinu než se síť naučila. Z tohoto důvodu bych doporučovala výkonnější počítač.

Jelikož jsou neuronové sítě dnes takovým fenoménem, doporučila bych na tuto problematiku zpracovat laboratorní úlohu, která by spočívala právě ve vyhledávání optimálního nastavení počtu vrstev a neuronů. Mohla by sloužit jak pro obory bezpečnostních technologií, kde by studenti dále testovali výstřely ze zbraní, tak pro obory informačních technologií, kde by se zabývali programováním v programovém prostředí Matlab.

## ZÁVĚR V ANGLIČTINĚ

The theme of this thesis was classification of audio signals using neural networks. This topic is very comprehensive, so, I decided to describe in the theoretical part of really important part. I was apprised of the numerator with classification, history and function of neural networks and problems with audio processing of signals.

The existing database that I had available from supervisor is not sufficient for this projekt, so we decided to practise of gunshots to create a new database. This work took place in the field and directly on the range Otrokovice. Gunfire was divided into 2 days. We had available the gun room. Arms were a total of 17 types pistols or revolvers and rifles. In particular, was emphasis on the creation of firearms, and the sound was the best recorded. To have achieved the highest quality sounds are used to capture high-quality microphone and units that were deployed within range.

The practical part was the first to create a database of sounds just mentioned and the other was to modify an existing program to signal classification using neural network.

The adjustment algorithm consisted in adding features to the extraction of characteristic feature of the signal. The extract was used for analysis of both the frequency and the time. Total area of each 2 characteristics were used. More in chapter 3.2. of these characteristics. After adjustment algorithm, which consisted in adding subprogrammes to extract the characteristic features of the signal, followed by experimentation and to try to find a network setup as the connection of layers and neurons, so that classification has been as successful as possible. The most plausible results in many instances confirmed the theory in the Experimental section, to say that two-layer network is showing good results. Depend on the success of a proper network settings, but also just the analysis, which extracted features from signal.

The only problem in this work was in experimental work. Specifically, the analysis time, while setting the number of neurons in the hidden layers of sound pressure. Setting the hidden layers of neurons more than [5;2], there was a break training because of the long training time. Time and many times exceeded 2 hours and percentage classification was small. It is possible that due to the performance of my PC or perhaps even both as seen in the Table 16 of the input layer, which contains a lot of input vector. I would recommend to reduce the reduction factor.

The outcome of this work is a modified program in Matlab 2009, created and modified learning set of sound. The task was to find a neural network parameters (number of neurons, layers, etc.) in order to practise what the maximum classification which in some instances succeeded.

This work includes experimental work, which detects what network settings are most appropriate for the characteristics of the signal with a database of patterns. As a continuation of the proposal I would recommend definitely more kinds of patterns in the database. The samples I tried to create in different environments. For example, a shooting range in an open environment. I also tried another analysis of a signal such as the spectrogram, which I use in this work not because I do not allow performance of my personal computers.

Parameters of my computers are Microsoft Windows 7 64Bit. Home Premium. Pentium Dual-Core T4300 2.10 GHz CPU, RAM 3.0 gigabytes. The parameters are mentioned because of some combination of network settings were so demanding that it lasted even an hour before the network is learned. For this reason, I would recommend more powerful computer.

Because neural networks are today such a phenomenon, I would recommend this issue to process laboratory tasks that would consist just to seek the optimal set of layers and neurons. It could also serve as security for the areas of technology, where students also tested the gunshot, and the fields of information technology, which would have dealt with programming in Matlab.

**SEZNAM POUŽITÉ LITERATURY**

- [1] KVASNIČKA, V. a kol. Úvod do teórie neurónových sietí. Bratislava: Iris, 1997. ISBN 80-88778-30-1. 262 s.
- [2] HLAVÁČ, Václav. Rozpoznávání s Markovskými modely [online]. Praha, 2010 [cit. 2011-05-05]. Dostupné z WWW: <<http://cmp.felk.cvut.cz/~hlavac/TeachPresCz/31Rozp/61MarkovianPR.pdf>>.
- [3] MAŘÍK, V.; ŠTEPÁNKOVÁ, O.; LAŽANSKÝ, J. a kolektiv. *Umělá inteligence 1*. [cit. 2011-04-17]. Academia, Praha, 1997.
- [4] MAŘÍK, V.; ŠTEPÁNKOVÁ, O.; LAŽANSKÝ, J. a kolektiv. *Umělá inteligence 4*. [cit. 2011-03-17]. Academia, Praha, 1997.
- [5] NEVŘINA, Pavel. *Analýza signálu a soustav*. 1. [s.l.] : [s.n.], 14.3.2002. 670 s. ISBN 80-7300-004-0.
- [6] ŠÍMA, J., NERUDA, R. Teoretické otázky neuronových sítí [online]. 1996 Dostupný z WWW: <<http://www2.cs.cas.cz/~sima/kniha.pdf>>.
- [7] KVASNIČKA, V., et al. Úvod do teórie neurónových sietí. Bratislava: Iris, 1997. 285 s. ISBN 80-88778-30-1.
- [8] KOSINA K.. Music Genre Recognition. Diplomová práce, Medientechnik und Design in Hagenberg , Hagenberg 2002. [cit. 2011-03-18].
- [9] KUBA, M. Neuronové sítě [online]. Brno: Masarykova univerzita, 1995 [cit. 2011-04-01]. Dostupné z WWW: <<http://tf.czu.cz/~votruba/WAN/Neuronove%20site.pdf>>.
- [10] HÁJEK, Ing. Petr; OLEJ, CSC., Ing. Vladimír. Úvod do umělé inteligence : Moderní přístupy. Pardubice: Univerzita Pardubice, 2010. 98 s. [cit. 2011-03-18].
- [11] KAČENKA, Petr. Neuronové sítě. Matematický korespondenční seminář [online]. 1998, 1, [cit. 2011-03-03]. Dostupný z WWW: <<http://mks.mff.cuni.cz/library/NeuronoveSitePK/NeuronoveSitePK.pdf>>.
- [12] KADLEC, CSC., DOC. ING. FRANTIŠEK. ZPRACOVÁNÍ AKUSTICKÝCH SIGNÁLŮ. [S.L.] : VYDAVATELSTVÍ ČVUT, 2002. 189 s.
- [13] PSUTKA J., MÜLLER L., MATOUŠEK J., RADOVÁ V.. MLUVÍME S POČÍTEČEM ČESKY. *Academia, Praha 2006*.
- [14] TAUFER, I.; DRÁBEK, O.; SEIDL, P. Umělé neuronové sítě – základy teorie a aplikace (5). *CHEMagazín*, XVI, 2006, č. 4., s. 29 – 31. ISSN 1210 – 7409.
- [15] Zpracování signálů a obrazů , Korelační metody, Petr Česák, 17.10.2002 online citace 4.5.2011

- [16] Horký L., Břinda K.: Neuronové sítě. Fakulta jaderná a fyzikálně inženýrská, Břehová 7, 115 19 Praha 1. Dostupný také z <http://fyzsem.fjfi.cvut.cz/2008-2009/Leto09/proc/neurony.pdf>
- [17] ZÁVODNÝ, Luděk; HANUS, Stanislav. Využití umělé neuronové sítě pro empirický model šíření signálu. *Vysoké učení technické v Brně*. 2004, 1636, s. 7. Dostupný také z WWW: <[http://dsp.vscht.cz/konference\\_matlab/matlab04/zavodny.pdf](http://dsp.vscht.cz/konference_matlab/matlab04/zavodny.pdf)>.
- [18] Algoritmy učení. *Katedra teorie obvodů* [online]. 2006, --, [cit. 2011-05-06]. Dostupný z WWW: <<http://amber.feld.cvut.cz/ssc/ssc-cv/bpg.pdf>>.
- [19] M. Novák a kolektiv: Umělé neuronové sítě, C. H. Beck, Praha, 1997, [cit. 2011-05-06].
- [20] Šidlovský, J.: Využitie neurónových sietí pre rozpoznávanie fliaš při recyklácii. Košice, 2005, Dostupný také z <http://neuron.tuke.sk/jaksa/theses/2005/Sidlovsky-Jaksa-MSc05-thesis.pdf>
- [21] *Význam jednotlivých bytov pro konkrétný WAVE soubor* [online]. 2010 [cit. 2011-06-02]. Dostupné z WWW: <<https://ccrma.stanford.edu/courses/422/projects/WaveFormat/wave-bytes.gif>>.
- [22] VLACHY, Vaclav. *Praxe zvukové techniky*. 2. aktualizované vydání. Praha : Nakladatelství Muzikus, 2000. ISBN 80-86253-05-8.
- [23] Digital signal processing: Wikipedia, the free encyclopedia [online]. [cit. 2007-10-30]. URL: [http://www.wikipedia.org=Digital signal processing:html](http://www.wikipedia.org=Digital%20signal%20processing:html) :
- [24] Ústav řízení systémů a spolehlivosti [online]. 2006, 04.06.2010 [cit. 2011-05-16]. Dostupné z: <[http://www.rss.tul.cz/download/cms/10\\_zakl\\_freq\\_anal.pdf](http://www.rss.tul.cz/download/cms/10_zakl_freq_anal.pdf)>.
- [25] Pavela, Bc. Radek. Využití neuronových sítí pro predikaci síťového provozu [online]. Brno, 2009. 83 s. Diplomová práce. Vysoké učení technické v Brně. [http://www.vutrb.cz/www\\_base/zav\\_prace\\_soubor\\_verejne.php?file\\_id=15208](http://www.vutrb.cz/www_base/zav_prace_soubor_verejne.php?file_id=15208).
- [26] HAYKIN, S. *Neural Networks: A Comprehensive Foundation*. 2nd edition, New Jersey, Prentice-Hall, Inc., 1999, 842s. ISBN 81-7808-300-0.
- [27] KOPEČEK, Ivan. *Zpracování digitalizovaného signálu*. Fakulta informatiky Masarykovy univerzity [online]. 2010, 4, [cit. 2010-04-28]. Dostupný z WWW: <[www.fi.muni.cz/~kopecek/upzr3.ppt](http://www.fi.muni.cz/~kopecek/upzr3.ppt)>.

- [28] *Saint-Gobain* [online]. 2002-2009 [cit. 2011-05-17]. Ecophon. Dostupné z WWW: <http://ecophon.ru/cz/Akustika/Uivatel/Akustikazvuk-e-a-slyitelnost/>.
- [29] DEMUTH, H., BEALE, M., HOGAN, M.: *Neural Network Toolbox™6 User's Guide* [online]. 2009 [cit. 2011-05-18]. Dostupné na internetu: <  
[http://www.mathworks.com/access/helpdesk/help/pdf\\_doc/nnet/nnet.pdf](http://www.mathworks.com/access/helpdesk/help/pdf_doc/nnet/nnet.pdf)>
- [30] N. Negroponte, *Digitální svět*. Praha: Management Press, 2001 - 207 s. ISBN 80-7261-046-5

**SEZNAM POUŽITÝCH SYMBOLŮ A ZKRATEK**

UNS	Umělé neuronové sítě
FT	„ <i>Fourier transform</i> “ (Fourierova transformace)
XOR	Kombinační logický obvod exklusive
FFNN	„ <i>Feedforward Neural Networks</i> “ (dopředné neuronové sítě)
FFT	„ <i>Fast Fourier Transform</i> “ (rychlá fourierova transformace)
DTF	„ <i>Discrete Fourier Transform</i> “ (diskrétní Fourierova transformace)
WAV(WAVE)	„ <i>Waveform audio format</i> “ (zvukový formát)
SCR	„ <i>Success Clasification Rate</i> “ (míra úspěšnosti neuronové sítě)
RCNN,RNN	„ <i>Recurrent Neural Network</i> “ (rekurentní neuronové sítě)
DSP	„ <i>Digital Signal Processing</i> “ (digitální signálový procesor)
MSE	„ <i>Mean Squared Error</i> “ (střední chyba)
SSE	„ <i>Sum Squared Error</i> “ (střední chyba)

**SEZNAM OBRÁZKŮ**

<i>Obrázek 1: Biologický neuron</i> .....	14
<i>Obrázek 2: Model umělého neuronu Zdroj: [10]</i> .....	16
<i>Obrázek 3: Sigmoidní funkce</i> .....	17
<i>Obrázek 4: Neuronové sítě. A - dopředná síť</i> .....	19
<i>Obrázek 5: Příklad dopředné neuronové sítě s jednou skrytou vrstvou[17]</i> .....	22
<i>Obrázek 6: Digitalizace analogových dat</i> .....	25
<i>Obrázek 7: Signál zaznamenaný mikrofonními jednotkou</i> .....	26
<i>Obrázek 8: Originální signál A - Průměrovaný signál B</i> .....	31
<i>Obrázek 9: Redukovaný signál pomocí autokorelační funkce</i> .....	32
<i>Obrázek 10: Originální signál A - Redukovaný akustický tlak B</i> .....	33
<i>Obrázek 11: Centrální momenty</i> .....	34
<i>Obrázek 12: Guncenter Otrokovice</i> .....	36
<i>Obrázek 13: Tvorba databáze</i> .....	36
<i>Obrázek 14: Propojení snímací techniky a PC</i> .....	41
<i>Obrázek 15: Mikrofon B&amp;K Type 4188</i> .....	42
<i>Obrázek 16: Předzesilovač B&amp;K Type 2671</i> .....	42
<i>Obrázek 17: Zdroj proudu</i> .....	43
<i>Obrázek 18: Mobilní jednotka pro sběr dat</i> .....	44
<i>Obrázek 19: Program Matlab</i> .....	46
<i>Obrázek 20: Grafická podoba trénovacího procesu</i> .....	47
<i>Obrázek 21: Závislost trénovací chyby na počtu epoch</i> .....	48

**SEZNAM TABULEK**

<i>Tabulka 1: Přehled vzorů v databázi 1</i> .....	38
<i>Tabulka 2: Přehled vzorů v databázi 2</i> .....	39
<i>Tabulka 3: Přehled vzorů v databázi 3</i> .....	40
<i>Tabulka 4: Přehled vzorů v databázi 4</i> .....	40
<i>Tabulka 5: Parametry mikrofonu B&amp;K Type 4188</i> .....	42
<i>Tabulka 6: Parametry předzesilovače B&amp;K Type 2671</i> .....	42
<i>Tabulka 7: Vstupní rozsahy jednotky USB-4716</i> .....	44
<i>Tabulka 8: Tabulka s legendou</i> .....	51
<i>Tabulka 9: Popis databáze výstřelů</i> .....	52
<i>Tabulka 10: Učící a trénovací cykly s použitím průměrového spektra</i> .....	53
<i>Tabulka 11: Učící a trénovací cykly s použitím průměrového spektra</i> .....	54
<i>Tabulka 12: Učící a trénovací cykly s použitím průměrového spektra</i> .....	55
<i>Tabulka 13: Učící a trénovací cykly s použitím průměrového spektra</i> .....	55
<i>Tabulka 14: Učící a trénovací cykly s použitím autokorelací fce</i> .....	56
<i>Tabulka 15: Učící a trénovací cykly s použitím autokorelací fce</i> .....	57
<i>Tabulka 16: Učící a trénovací cykly s použitím</i> .....	57
<i>Tabulka 17: Učící a trénovací cykly s použitím autokorelací fce</i> .....	58
<i>Tabulka 18: Učící a trénovací cykly s použitím akustického tlaku</i> .....	59
<i>Tabulka 19: Učící a trénovací cykly s použitím akustického tlaku</i> .....	60
<i>Tabulka 20: Učící a trénovací cykly s použitím akustického tlaku</i> .....	61
<i>Tabulka 21: Učící a trénovací cykly s použitím akustického tlaku</i> .....	61
<i>Tabulka 22: Učící a trénovací cykly s použitím centrálních momentů</i> .....	62
<i>Tabulka 23: Učící a trénovací cykly s použitím centrálních momentů</i> .....	63
<i>Tabulka 24: Učící a trénovací cykly s použitím</i> .....	63
<i>Tabulka 25: Učící a trénovací cykly s použitím centrálních momentů</i> .....	64

---

## SEZNAM PŘÍLOH

*Příloha P I* .....*Obsah CD*

*Příloha P II* .....*Jádro algoritmu*

## **PŘÍLOHA P I: OBSAH CD**

### 1. Diplomová práce:

Klasifikace audio signálu pomocí neuronových sítí. docx

Klasifikace audio signálu pomocí neuronových sítí. pdf

### 2. Data:

a. Databáze\_1A

b. Databáze\_1B

c. Databáze\_2A

d. Databáze\_2B

e. Databáze\_3A

f. Databáze\_3B

g. Databáze\_4A

h. Databáze\_4B

### 3. Program a podprogramy na klasifikaci:

a. Start.m

b. Fce\_get\_audiodata\_from\_dat.m

c. Fce\_get\_spectrum\_from\_dat.m

d. Fce\_klouz\_prumer.m

e. Fce\_reduction.m

f. Fce\_teziste.m

g. Fce\_window\_function.m

h. Feature1.m (autokorelační funkce)

i. Feature2.m (redukovaný akustický tlak)

j. Feature3.m (průměrované spektrum)

k. Feature4.m (centrální momenty)

l. Mojedata.mat

## PŘÍLOHA P II: JÁDRO ALGORITMU

```
clc;
tic;

TrainDIR = 'data'; %adresar obsahujici trénovací množinu dat

%% Definovani feature vektorů (na extrakci charakteristických
rysů ze signálu
feature = 3;

Posun = 0.1; %odstraneni zacatku zvuku
Delka = 1.4; %delka signalu braneho k analyze - v casove
oblasti

%% FEATURE 1 - autokorelační funkce
if(feature==1)

    NoPointsFFT = 16384;
    MinFREQ = 100;
    MaxFREQ = 5000;
    Reduction_factor = 5;
    NoPoinstOf_MA = 11;

end

%% FEATURE 2 - redukovany akusticky tlak
if(feature==2)

    Reduction_factor = 40;

end

%% FEATURE 3 - průměrované spektrum
if(feature==3)

    NoPointsFFT = 16384;
    MinFREQ = 100;
```

```

MaxFREQ = 3000;
Krok = 10; %prumerovat budu po N bodech

end

%% FEATURE 4 - centrální momenty
if(feature==4)

    Pocet_M = 4; %počet momentu

end

%% Příprava dat pro neuronovou sit
disp('Processing data ...');
k=dir(TrainDIR);
pocet_polozek = length(k);

if (pocet_polozek==0)
    return;
end

    skupina='';
% vytvoří se seznam adresářů
for i=1:pocet_polozek
    if k(i).isdir==true
        if
((strcmp(k(i).name, '.')==0)&&(strcmp(k(i).name, '..')==0))
            skupina=strvcat(skupina, k(i).name);
        end
    end
end

% počet adresářů
adr=size(skupina,1);

% celkový počet zvukových souborů
souboru=0;

% příprava výstupního vektoru pro trénovací množinu -
výstupem je daná třída
Tc=[];
mag=[];

for a=1:adr
    newpath=[TrainDIR '\\' deblank(skupina(a,:)) '\\'];
    % načtení všech zvukových souborů z daného adresáře
    k=dir([newpath '*.dat']);
    pocet_polozek = length(k);

```

```

%           set(handles.txtStatus,'String',['Processing
files from database (' int2str(a) '/' int2str(adr) ')...']);

    for i=1:pocet_polozek

        if k(i).isdir==false
            souboru=souboru+1;
            Tc=[Tc a];
            sndfile=k(i).name;

            if(feature==1)
                FV=feature1([newpath     sndfile],     Posun,
NoPointsFFT,     MinFREQ,     MaxFREQ,     Reduction_factor,
NoPoinstOf_MA);
            end

            if(feature==2)
                FV=feature2([newpath     sndfile],     Posun,
Delka, Reduction_factor);
            end

            if(feature==3)
                FV=feature3([newpath     sndfile]     ,Posun,
NoPointsFFT, MinFREQ, MaxFREQ, Krok);
            end

            if(feature==4)
                FV=feature4([newpath     sndfile],     Posun,
Delka, Pocet_M);
            end
            mag(souboru,:) = FV;
        end
    end

    end

% vstupní vektor
P = mag';
% výstupní vektor
T = ind2vec(Tc);

% Normalize Inputs and Targets
[normInput,ps] = mapminmax(P,0,1);
[normTarget,ts] = mapminmax(T,0,1);

```

```

save mojedata

load mojedata

%% Neuronová síť - train
Neurons_HL = '5 2';
TransFcns = { 'logsig' 'logsig' 'logsig'}; %Log-sigmoid
transfer function
BTF = 'trainlm'; %Levenberg-Marquardt Back
Propagation, metoda učení
BLF = 'learngdm'; %Gradient descent with momentum
weight/bias learning fiction, určení vah
PF = 'sse'; %Sum squared error performance
fiction, součet chyb

Layers = eval(['[ ' Neurons_HL ' ' num2str(adr) ']'']);

net = newff(minmax(normInput),Layers,TransFcns,BTF,BLF,PF);

% net.divideparam.trainInd=1:30;
% net.divideparam.valInd=31:34;
% net.divideparam.testInd=35:40;

%Inicializace sítě
% net.inputWeights{1,1}.initFcn = 'initzero';
% net = init(net);

testPercent = 0.5; % 50% vzorů je určeno na testovani
validatePercent = 0.1; % 10% na validaci
[trainSamples,testSamples, validateSamples] =
dividevec(normInput,normTarget,testPercent,validatePercent);

%ovlivneni zobrazeni v okne a v prikaz.radku
net.trainParam.showCommandLine = 0;
net.trainParam.showWindow = 1;

net.trainParam.mu = 0.001 ;
net.trainParam.mu_dec = 0.1 ;
net.trainParam.mu_inc = 10 ;
% net.trainParam.show = 5 ;
net.trainParam.epochs = 200 ;

```

```

disp('Training process ...');
%Trenovani site
[net,tr] = train(net,trainSamples.P,trainSamples.T,[],
[],validateSamples,testSamples);

[normTrainOutput,Pf,Af,E,trainPerf] =
sim(net,trainSamples.P,[], ...
[],trainSamples.T);
[normValidateOutput,Pf,Af,E,validatePerf] =
sim(net,validateSamples.P, ...
[],[],validateSamples.T);
[normTestOutput,Pf,Af,E,testPerf] = sim(net,testSamples.P,[],
...
[],testSamples.T);
%
%
trainOutput = mapminmax('reverse',normTrainOutput,ts);
validateOutput = mapminmax('reverse',normValidateOutput,ts);
testOutput = mapminmax('reverse',normTestOutput,ts);

[values,testSamplesIndex]=max(testSamples.T);
[values,testOutputIndex]=max(normTestOutput);

[values,trainSamplesIndex]=max(trainSamples.T);
[values,trainOutputIndex]=max(normTrainOutput);

nTest = size(normTestOutput);
nTest = nTest(2);

nTrain = size(normTrainOutput);
nTrain = nTrain(2);

scrTest = sum(testSamplesIndex==testOutputIndex) / nTest *
100;
scrTrain = sum(trainSamplesIndex==trainOutputIndex) / nTrain
* 100;

t=toc;

disp(['Total time: ' num2str(t,3) ' s']);
disp(['Train succes rate: ' num2str(scrTrain,3) ' %']);
disp(['Test succes rate: ' num2str(scrTest,3) ' %']);

```