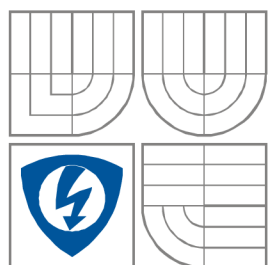


VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH
TECHNOLOGIÍ

ÚSTAV BIOMEDICÍNSKÉHO INŽENÝRSTVÍ

FACULTY OF ELECTRICAL ENGINEERING AND
COMMUNICATION

DEPARTMENT OF BIOMEDICAL ENGINEERING

KLASIFIKACE SRDEČNÍCH CYKLŮ Z VÍCE SVODOVÉHO EKG POMOCÍ METODY HLAVNÍCH KOMPONENT

CLASSIFICATION OF HEART BEATS FROM MULTILEAD ECG USING PRINCIPAL
COMPONENT ANALYSIS

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

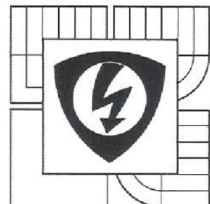
Bc. Milan Vlček

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Marina Ronzhina

BRNO, 2013



VYSOKÉ UČENÍ
TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

Ústav biomedicínského inženýrství

Diplomová práce

magisterský navazující studijní obor
Biomedicínské inženýrství a bioinformatika

Student: Bc. Milan Vlček
Ročník: 2

ID: 119760
Akademický rok: 2012/13

NÁZEV TÉMATU:

Klasifikace srdečních cyklů z více svodového EKG pomocí metody hlavních komponent

POKYNY PRO VYPRACOVÁNÍ:

1) Seznamte se s základními principy metody hlavních komponent (angl. principal component analysis, PCA). Zaměřte se na její využití při klasifikaci srdečních cyklů. Proveďte literární rešerši v této oblasti. 2) Sestavte databázi normálních a patologických srdečních cyklů za použitím EKG z veřejných databází nebo signálů dostupných na ÚBMI. Použijte alespoň 3 svody EKG. 3) Aplikujte PCA na srdeční cykly z navržené databázi. Z vypočtených komponentů vyberte nejvíce informativní. Využijte programové prostředí MATLAB. 4) Navrhněte klasifikátor srdečních cyklů (např. vícevrstvou umělou neuronovou síť) na základě vybraných PCA parametrů. 5) Proveďte hodnocení výsledků klasifikace. Porovnejte dosažené výsledky s dostupnými údaji z literatury. 6) Proveďte diskusi získaných výsledků.

DOPORUČENÁ LITERATURA:

[1] CEYLAN, Rahime a ÖZBAY, Yüksel. Comparison of FCM, PCA and WT techniques for classification ECG arrhythmias using artificial neural network. Expert Systems with Applications. 2007, s. 286-295.
[2] CLIFFORD, Gary D, AZUAJE, Francisco a MCSHARRY, Patrick. Advanced methods and tools for ECG data analysis. Boston: Artech House, 2006. ISBN 1-58053-966-1.

Termín zadání: 11.2.2013

Termín odevzdání: 24.5.2013

Vedoucí práce: Ing. Marina Ronzhina
Konzultanti diplomové práce:



prof. Ing. Ivo Provazník, Ph.D.
předseda oborové rady

UPOZORNĚNÍ:

Autor diplomové práce nesmí při vytváření diplomové práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

ABSTRAKT

Cílem této diplomové práce je seznámit se s metodou hlavních komponent. Dále se zaměřit na její využití při zpracování srdečních cyklů. Tato metoda umožňuje zredukovat množství dat beze ztráty užitečných informací, což je důvodem proč se v posledních letech často používá pro zpracování dat a pro jejich následnou klasifikaci, kterou se tato práce také zabývá. Data byla získána ústavem biomedicínského inženýrství na FEKT VUT v Brně. Jejich následná analýza byla provedena v programu Matlab.

KLÍČOVÁ SLOVA

Elektrokardiogram, srdeční cyklus, zpracování EKG, analýza hlavních komponent, umělé neuronové sítě

ABSTRACT

The resume of this master's thesis is to introduce reader into principal component analysis (PCA), namely, the use of PCA for analysis of ECG. This method allows to reduce quantity of the data without loss of useful information. That is why PCA is widespread for preprocessing of the data for further classification, which this thesis also deals. Data available at the Department of Biomedical Engineering at the University of Technology in Brno were used in this work. All the methods were realized using Matlab.

KEYWORDS

Electrocardiogram, heart beat, ECG processing, principal component analysis, artificial neural networks

VLČEK, M. *Klasifikace srdečních cyklů z více svodového EKG pomocí metody hlavních komponent*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2013. 66 s. Vedoucí diplomové práce Ing. Marina Ronzhina.

PROHLÁŠENÍ

Prohlašuji, že svou diplomovou práci na téma Klasifikace srdečních cyklů z více svodového EKG pomocí metody hlavních komponent jsem vypracoval samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené diplomové práce dále prohlašuji, že v souvislosti s vytvořením této diplomové práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědom následků porušení ustanovení § 11 a následujících zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

V Brně dne

.....

(podpis autora)

PODĚKOVÁNÍ

Děkuji vedoucí mé diplomové práce Ing. Marině Ronzhině za účinnou metodickou, pedagogickou a odbornou pomoc a další cenné rady při zpracování mé diplomové práce.

V Brně dne

.....

(podpis autora)

OBSAH

Seznam obrázků	8
Seznam tabulek	10
Úvod	11
1 SRDEČNÍ AKTIVITA A JEJÍ SNÍMÁNÍ	12
1.1 Elektrická aktivita srdce.....	12
1.2 Elektrokardiografie	13
1.2.1 Einthovenovy svody	13
1.2.2 Goldbergerovy svody.....	13
1.2.3 Hrudní svody.....	14
1.2.4 Ortogonální svody.....	15
1.3 EKG signál.....	15
2 ANALÝZA HLAVNÍCH KOMPONENT	17
2.1 Princip analýzy hlavních komponent.....	17
2.2 Využití PCA při zpracování EKG	21
3 UMĚLÉ NEURONOVÉ SÍTĚ	23
3.1 Perceptron	23
3.2 Topologie sítí	26
3.3 Učení neuronových sítí.....	27
4 KLASIFIKACE EXPERIMENTÁLNÍCH DAT	29
4.1 Popis experimentálních dat.....	29
4.2 Předzpracování dat.....	29
4.3 Zpracování experimentálních dat pomocí PCA.....	30
4.3.1 Aplikování PCA na sto cyklů z jednoho svodu	33
4.3.2 Zpracování tří svodů pomocí PCA	35
4.3.3 Využití PCA k redukci dat v rámci jednoho segmentu	36
4.4 Návrh klasifikátorů	42
4.5 Aplikace pro zpracování dat pomocí PCA a klasifikaci pomocí UNS ...	46
4.5.1 Aplikace s použitím PCA na 100 cyklů.....	46
4.5.2 Aplikace s použitím PCA na všechny tři svody	51
4.5.3 Aplikace s využitím redukce dat v rámci jednoho cyklu.....	55

4.6	Výsledky klasifikace QRST segmentů	59
5	ZÁVĚR	63
	Literatura	65

SEZNAM OBRÁZKŮ

Obr. 1.1:	Převodní srdeční systém [4].....	12
Obr. 1.2:	Einthovenovy svody.....	13
Obr. 1.3:	Goldbergerovy svody.....	14
Obr. 1.4:	Rozmístění hrudních elektrod[3]	14
Obr. 1.5:	Frankův systém – rozmístění elektrod [5].	15
Obr. 2.1:	Schéma výpočtu hlavních komponent.	19
Obr. 2.2:	Ukázka-scree plot.....	20
Obr. 3.1:	Jednoduchý perceptron	23
Obr. 3.2:	Rozdělení do dvou tříd s využitím jednoho neuronu	24
Obr. 3.3:	Skoková aktivační funkce	25
Obr. 3.4:	Sigmoidní funkce	25
Obr. 3.5:	Vícevrstvá neuronová síť s jednou skrytou vrstvou[25].....	26
Obr. 4.1:	Schéma zpracování dat pomocí PCA: výběr 1 repre. segmentu ze 100	31
Obr. 4.2:	Schéma zpracování dat pomocí PCA: výběr 1 repre. segmentu ze 3 svodů	31
Obr. 4.3:	Schéma zpracování dat pomocí PCA: redukce v rámci jednoho segmentu.	32
Obr. 4.4:	Sto QRST segmentů vybraných z kontrolní fáze.....	33
Obr. 4.5:	Scree plot, sto cyklů.....	34
Obr. 4.6:	Ukázka hlavních komponent (1k-57%,2k.90%,3k-97% variability).....	34
Obr. 4.7:	Scree plot, tři svody	35
Obr. 4.8:	Signály se tří svodů (X, Y, Z) a výstup PCA.....	36
Obr. 4.9:	Scree plot, redukce dat.....	37
Obr. 4.10:	Násobení vstupního cyklu se zredukovanou maticí koeficientů.....	38
Obr. 4.11:	Získání rekonstruovaného cyklu pomocí zredukované matice koeficientů. 38	
Obr. 4.12:	Původní a rekonstruovaný cyklus (2 komponenty, korelace=0,73).....	39
Obr. 4.13:	Původní a rekonstruovaný cyklus (20 komponent, korelace=0,99).....	40
Obr. 4.14:	Cyklus rekonstruovaný ze všech komponent (korelace=1)	40
Obr. 4.15:	Prvních 20 komponent vstupního cyklu	41
Obr. 4.16:	Náhodně vybrané normální a ischemické cykly (5000+5000), svod X.....	47
Obr. 4.17:	Vybrané ukázkové cykly: normální a ischemické	47
Obr. 4.18:	Scree ploty pro normální (vlevo) a ischemická data (vpravo).....	48

Obr. 4.19: První dvě komponenty pro vybrané cykly-normální(z) a ischemické(č).....	48
Obr. 4.20: První dvě komponenty všech cyklů: normálních (z) a ischemických (č)	49
Obr. 4.21: Aplikace Jeden ze sta	49
Obr. 4.22: Učení neuronové sítě.....	50
Obr. 4.23: Výsledek klasifikace	51
Obr. 4.24: Náhodně vybrané signály, svod X, Y a Z (3x50)	52
Obr. 4.25: Scree plot pro normální fázi.....	52
Obr. 4.26: Tři vstupní signály (svody X, Y, Z) a výsledná komponenta	53
Obr. 4.27: Výsledné hlavní komponenty ze tří svodů.....	54
Obr. 4.28: Aplikace Tři svody.....	54
Obr. 4.29 Náhodně vybrané cykly pro obě fáze (50+50), svod Z.....	56
Obr. 4.30: Jeden vybraný cyklus z každé fáze, svod Z.....	56
Obr. 4.31: Prvních 20 komponent vybraných cyklů	57
Obr. 4.32: Rekonstruovaný cykly pomocí 20-ti komponent.....	57
Obr. 4.33: Prvních 20 komponent vybraných cyklů	58
Obr. 4.34: Aplikace redukce	58

SEZNAM TABULEK

Tab. 2.1:	Přehled úspěšnosti klasifikace při použití PCA k výpočtu příznaků	21
Tab. 2.2:	Přehled používaných metod	22
Tab. 4.1:	Ukázka souboru dat s QRST segmenty (výber z jednoho experimentu)	30
Tab. 4.2:	Úspěšnost klasifikace při různém počtu neuronů 1	43
Tab. 4.3:	Úspěšnost klasifikace při různém počtu epoch 1	43
Tab. 4.4:	Úspěšnost klasifikace při různém počtu neuronů 2	44
Tab. 4.5:	Úspěšnost klasifikace při různém počtu epoch 2	44
Tab. 4.6:	Úspěšnost klasifikace při různém počtu skrytých vrstev 1	45
Tab. 4.7:	Úspěšnost klasifikace při různém počtu neuronů 3	45
Tab. 4.8:	Úspěšnost klasifikace při různém počtu epoch 3	45
Tab. 4.9:	Úspěšnost klasifikace při různém počtu skrytých vrstev 2	46
Tab. 4.10:	Úspěšnost klasifikace pro 1 ze 100, svod X	59
Tab. 4.11:	Úspěšnost klasifikace pro 1 ze 100, svod Y	60
Tab. 4.12:	Úspěšnost klasifikace pro 1 ze 100, svod Z	60
Tab. 4.13:	Úspěšnost klasifikace, tři svody	60
Tab. 4.14:	Úspěšnost klasifikace, svod X	61
Tab. 4.15:	Úspěšnost klasifikace, svod Y	61
Tab. 4.16:	Úspěšnost klasifikace, svod Z	61

ÚVOD

Základní metodou pro vyšetření srdce je elektrokardiografie, která snímá elektrickou aktivitu srdce. Výsledkem je elektrokardiogram (EKG), obsahující nasnímané průběhy srdečních cyklů, podle kterého mohou být rozpoznány patologické stavy, mezi které patří například ischemie. Ischemie může být v EKG rozpoznána podle určitých změn, především podle QRS komplexu a tvaru S-T úseku. V současné době jsou elektrokardiogramy získávány převážně v digitální formě, což umožňuje jejich zpracování a analyzování pomocí počítačové techniky různými metodami. Příkladem může být analýza variability srdečního rytmu, využití kumulantů vyšších řádů nebo metoda hlavních komponent.

Metoda hlavních komponent je původem statistická metoda, která se ale postupem času začala využívat také při zpracování signálů. Jejím základním principem je transformace dat do jiných souřadnic. Výstupní data jsou pak dána lineární kombinací původních a jsou rozdělena do tzv. hlavních komponent. Každá komponenta obsáhne určitou část informace z původních dat. Díky této vlastnosti se metoda dá využít při analýze biomedicínských dat, jako jsou srdeční cykly. Může být využita ke snížení počtu srdečních cyklů. V tomto případě jsou na vstup dávány desítky srdečních cyklů. Výsledkem je jeden srdeční cyklus reprezentovaný jednou nebo více hlavními komponentami, který obsahuje nejpodstatnější informace.

Srdeční cykly analyzované metodou hlavních komponent mohou být klasifikovány pomocí některého z vhodných algoritmů. Mezi nejpoužívanější patří umělá neuronová síť. Umělá neuronová síť byla použita i v této práci, konkrétně se jedná o vícevrstvou dopřednou neuronovou síť.

Práce je rozdělena na několik částí. Nejprve na teoretickou, kde jsou popsány nejdůležitější věci, které je třeba vědět o srdci a o zpracování EKG metodou hlavních komponent. Dále jsou uvedeny základní typy neuronových sítí a jejich použití. V praktické části je popsán postup zpracování dat pomocí vytvořených aplikací, graficky jsou zobrazeny výsledky, které jsou okomentovány. Výsledky celé práce jsou shrnuty v závěru.

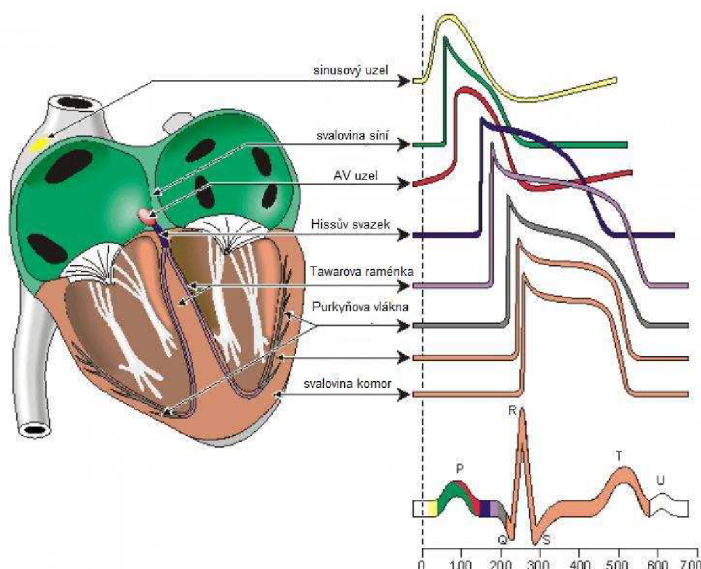
1 SRDEČNÍ AKTIVITA A JEJÍ SNÍMÁNÍ

Mechanická činnost srdce je řízena elektricky. Z těchto elektrických impulzů získává srdce přesné informace o tom, jakou mechanickou činnost má právě provést (stažení, relaxace apod.). Snímáním elektrických impulzů lze získat komplexní informace o elektrické aktivitě srdce.

1.1 Elektrická aktivita srdce

Srdce je řízeno pravidelnými elektrickými impulzy, které vznikají v sinoatriálním (SA) uzlu. Ten se nachází v pravé síni, konkrétně v její horní části. SA uzel je tedy hlavní srdeční stimulátor, který určuje srdeční rytmus. Vzruch dále pokračuje do atrioventrikulárního (AV) uzlu, který se nachází ve spodní části pravé síně. V AV uzlu také mohou vznikat elektrické impulzy, ale mají výrazně nižší frekvenci než impulzy vznikající v SA uzlu. Proto se tyto impulzy běžně na srdeční činnosti nijak neprojeví. AV uzel plynule přechází na Hisův svazek, který slouží k propojení síní a komor, na Hisův svazek, navazují Tawarova raménka (pravé a levé), ze kterých vystupují Purkyňova vlákna [1].

Výše uvedené prvky (SA uzel, AV uzel, Hisův svazek, Tawarova raménka, Purkyňova vlákna) tvoří dohromady *převodní srdeční systém* (obr. 1.1) [1]. V každém bodě (SA uzlu, AV uzlu apod.), kterým signál prochází, má jinou amplitudu a jiný tvar. Složením těchto křivek (tzv. akčních potenciálů) lze získat výsledný EKG signál.



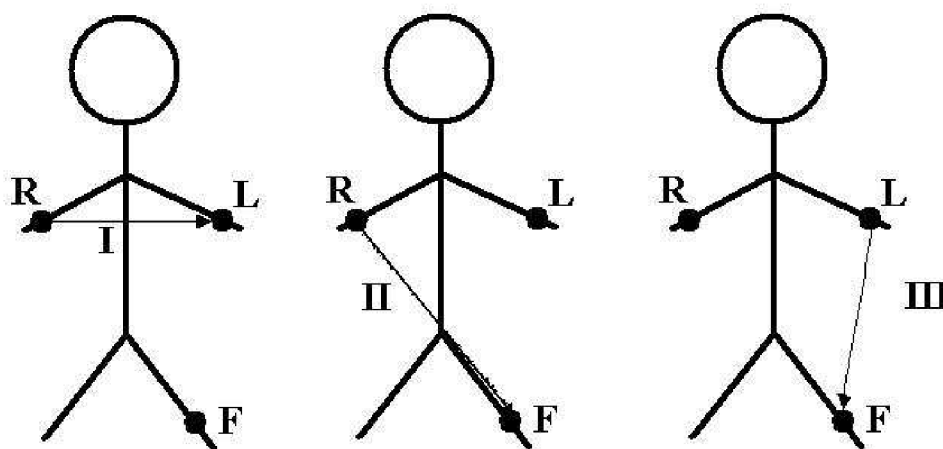
Obr. 1.1: Převodní srdeční systém [4]

1.2 Elektrokardiografie

Elektrokardiografie (EKG) je metoda, pomocí které lze měřit elektrickou aktivitu srdce. Velkou výhodou je, že měření lze provádět neinvazivně. Invazivní měření, tedy snímání přímo z povrchu srdce se provádí pouze ve výjimečných případech. Obvykle se měření provádí v klidovém stavu. Při standardním měření je na tělo pacienta upevněno deset elektrod (čtyři na končetiny a šest na hrudník), pomocí kterých je získán dvanáctisvodový elektrokardiogram. Jedná se o dvanáct EKG křivek, získaných měřením napětí mezi různými body. Existují tři základní typy svodů: Einthovenovy, Goldbergerovy a hrudní svody. Speciálním případem jsou ortogonální svody [2].

1.2.1 Einthovenovy svody

Tyto svody jsou také nazývány jako končetinové. Jsou celkem tři a označují se římskými číslicemi I, II a III. Svod I udává napětí mezi levou a pravou rukou. Napětí mezi pravou rukou a levou nohou je označeno jako II. A poslední svod III udává napětí mezi levou rukou a levou nohou. Einthovenovy svody jsou na obrázku 1.2 [2].

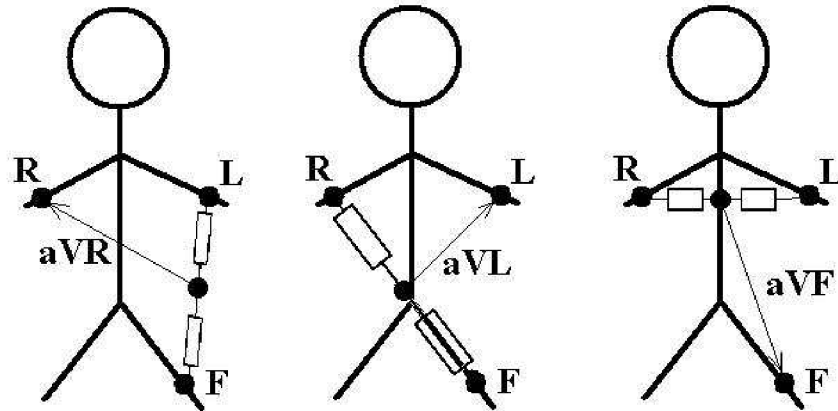


Obr. 1.2: Einthovenovy svody I, II, III

1.2.2 Goldbergerovy svody

Tyto svody využívají vždy všech tří měřících končetinových elektrod. Zjišťují napětí mezi jednou končetinou a průměrem zbylých dvou končetin. Mezi elektrodami jsou umístěny rezistory o stejné hodnotě, obvyklá hodnota je 5 kiloohmů. Označení těchto svodů začíná písmenem *a*, což značí zkratku slova augmentované, což znamená

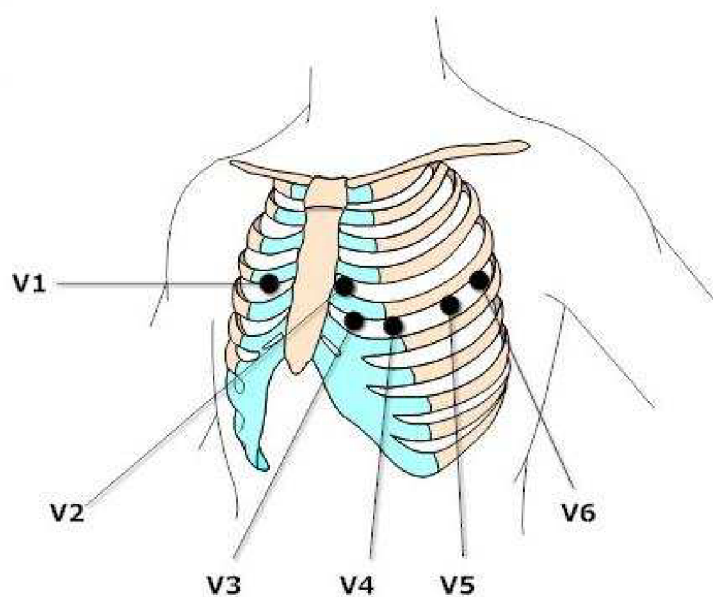
navýšené nebo zesílené. Goldbergerovy svody jsou tři: aVL, aVR, aVF, viz obr. 1.3 [2].



Obr. 1.3: Goldbergerovy svody: aVR, aVL, aVF

1.2.3 Hrudní svody

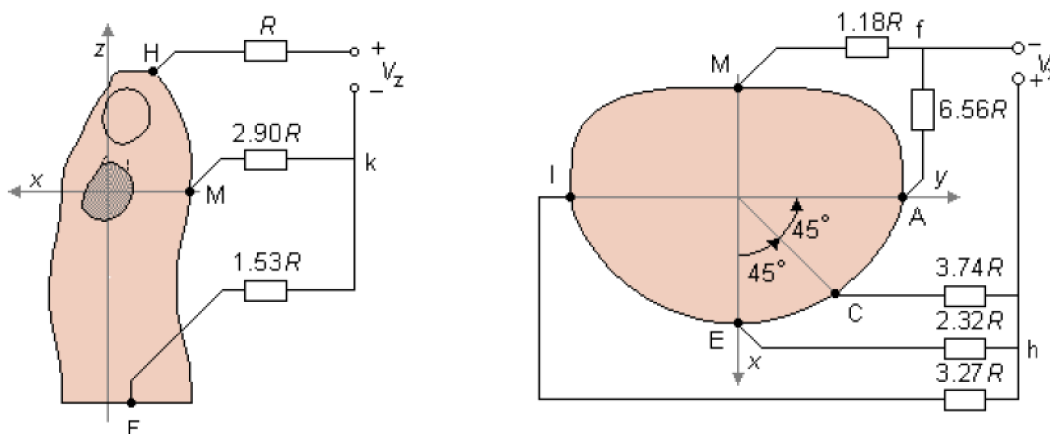
Hrudních svodů je celkem šest a jsou získány pomocí šesti elektrod rozmístěných na hrudníku dle určitého standardu (obr. 1.4). Zde se využívá speciálního zapojení, díky kterému je vytvořena tzv. *indiferentní* elektroda, která má nulový potenciál. Právě mezi touto elektrodou (s nulovým potenciálem) a mezi každou hrudní elektrodou je měřeno napětí (šest elektrod = šest svodů). Tyto svody se označují V1 – V6 [2].



Obr. 1.4: Rozmístění hrudních elektrod [3]

1.2.4 Ortogonální svody

Tyto svody dokážou vytvořit lepší prostorovou představivost o elektrické aktivitě. Ortogonální svody jsou založeny na tom, že z několika elektrod se vytvoří jeden svod. Napětí snímané několika elektrodami je převedeno na odporovou síť. Výstupem jsou tři svody označené U_x , U_y , U_z . Bylo vytvořeno několik systémů, založených na této myšlence, nejznámější je *Frankův systém* (obr. 1.5) [5].



Obr. 1.5: Frankův systém – rozmístění elektrod [5]

1.3 EKG signál

Každá část EKG křivky reprezentuje určitou mechanickou činnost srdce. Základními prvky EKG signálu jsou: P vlna, QRS komplex a T vlna. Na základě jejich amplitud a doby trvání lze určit informace o zdravotním stavu srdce. Sledovaných morfologických parametrů v EKG signálu je však více[1]:

- *P-vlna* vzniká při činnosti SA uzlu a znázorňuje depolarizaci síní. Maximální doba trvání je 300 ms a amplituda do 0,3 mV.
- *P-R interval* začíná depolarizací síní a končí depolarizací komor. Běžná doba trvání se pohybuje mezi 120 a 200 ms.
- *Q-vlna* je specifická tím, že má negativní výchylku. Standardně doba trvání nepřekračuje 30 ms.
- *R-vlna* následuje po vlně Q. Doba trvání je maximálně 100 ms a její amplituda může dosahovat až několik jednotek milivoltů.

- *S-vlna* má podobně jako *Q-vlna* negativní výchylku. Amplituda nepřekračuje 0,8 mV a doba trvání 50 ms.
- *QRS komplex* je nejvýznamnější částí EKG signálu. V této části signálu dochází k repolarizaci síní a depolarizaci komor. Protože depolarizace komor je doprovázena výrazně vyšším napětím než repolarizace síní, v křivce je patrná pouze depolarizace komor. Doba trvání se pohybuje mezi 60 a 120 ms.
- *S-T interval* reprezentuje čas, po který jsou komory aktivovány.
- *T-vlna* trvá 100 až 250 ms a reprezentuje repolarizaci komor.

2 ANALÝZA HLAVNÍCH KOMPONENT

Analýza hlavních komponent neboli PCA (z angl. Principal Component Analysis) je statistická metoda, jejímž účelem transformovat původní proměnné, které spolu korelují na jiné proměnné, které spolu korelovat nebudou. PCA může být využita v analýze biomedicínských dat ke kompresi biologických signálů, případně přímo k samotné klasifikaci. Výstupních proměnných lze vyprat pouze několik, a přesto vystihnou téměř veškerou informaci z původních proměnných[6].

2.1 Princip analýzy hlavních komponent

Prvním krokem je výpočet *kovarianční* nebo *korelační matice* pro data, která budou analyzována.

Kovarianční matice lze použít v případě, kdy vstupní data mají stejný význam, stejné jednotky apod. Kovarianční matice popisuje vzájemnou závislost proměnných. Matice je čtvercová, symetrická podle diagonály, na diagonále je uveden rozptyl příslušných proměnných a mimo diagonálu jsou odpovídající hodnoty kovariance. Kovariance mezi dvěma proměnnými je dána vztahem[10]:

$$\text{cov}_{xy} = \sum_{k=0}^{N-1} [x(k) - \mu_x][y(k) - \mu_y] \quad (1)$$

kde cov_{xy} je kovariance mezi proměnnými x a y , μ je střední hodnota, $x(k)$ je k -tý vzorek proměnné x , $y(k)$ je k -tý vzorek proměnné y a N je počet vzorků (délka) proměnných.

Korelační matice je třeba použít, pokud má dojít ke sledování vztahu mezi proměnnými o nesrovnatelných jednotkách (např. věku a krevního tlaku). I korelační matice je čtvercová a symetrická podle diagonály, ovšem na diagonále je číslo 1. To znamená, že každá proměnná maximálně koreluje sama se sebou, jedná se o standardizovaný rozptyl. Korelace je tedy kovariance na standardizovaných datech[10]. Ostatní hodnoty v matici udávají korelaci mezi proměnnými. Korelaci mezi dvěma proměnnými je možné spočítat podle vztahu[10]:

$$\text{cor}_{xy} = \frac{\text{cov}(x, y)}{\sqrt{\sigma_x^2 * \sigma_y^2}} \quad (2)$$

kde cor_{xy} je korelace mezi proměnnými x a y , $\text{cov}(x, y)$ je kovariance mezi proměnnými x a y , σ_x^2 je rozptyl proměnné x a σ_y^2 je rozptyl proměnné y .

Vlastní čísla

Pokud je známa korelační nebo kovarianční matice, nastává další krok, kterým je výpočet vlastních čísel podle vztahu[10]:

$$|A - \lambda_i I| = 0 \quad (3)$$

kde A je matice, pro kterou jsou počítána vlastní čísla (v tomto případě korelační nebo kovarianční), λ_i je i -té vlastní číslo a I je jednotková matice.

Pro kovarianční matici o velikosti 2×2 s prvky A, B, C, D bude výpočet následující[10]:

$$\left[\begin{array}{cc} A & B \\ C & D \end{array} \right] - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 0 \quad (4)$$

kde A, B, C, D jsou hodnoty kovarianční matice dat a λ je vlastní číslo.

Po vyřešení vztahu (4) budou parametru λ odpovídat dvě hodnoty, jsou tedy získána dvě vlastní čísla: λ_1 a λ_2 . Pomocí vlastních čísel se stanovuje, kolik variability vyčerpají vypočítané komponenty.

Vlastní vektory

Vlastní čísla jsou nutná k výpočtu vlastních vektorů, které jsou také nazývány vektory koeficientů. Vlastní vektor se počítá podle vztahu[10]:

$$(A - \lambda_i I) * u_i = 0 \quad (5)$$

kde A je kovarianční matice, λ_i je i -té vlastní číslo, I je jednotková matice a u_i je i -tý vlastní vektor.

V návaznosti na příkladu pro vlastní čísla pro matici o rozměrech 2×2 je postup následující[10]:

$$\left(\left[\begin{array}{cc} A & B \\ C & D \end{array} \right] - \lambda_i \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = 0 \quad (6)$$

kde A, B, C, D jsou hodnoty kovarianční matice dat, λ_i je i -té vlastní číslo, u_1 je první prvek vlastního vektoru, u_2 je druhý prvek vlastního vektoru.

Při dosazení vlastního čísla λ_1 bude vypočten vlastní vektor odpovídající prvnímu vlastnímu číslu, po dosazení vlastního čísla λ_2 dojde k výpočtu vlastního vektoru

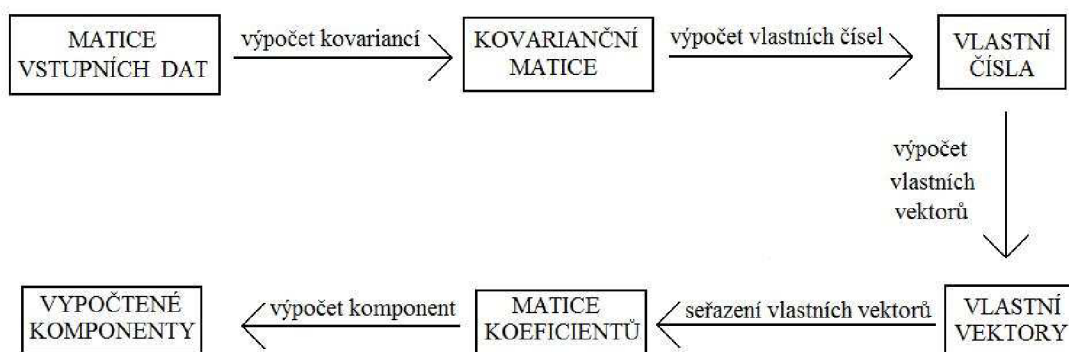
odpovídajícího druhému vlastnímu číslu. Vypočítané vlastní vektory se následně seskládají do jedné matice vlastních vektorů, nazývané také maticí koeficientů, přičemž sloupce odpovídají jednotlivým vektorům. Výstupní proměnné se označují jako **hlavní komponenty**, které jsou dány součinem původních proměnných s příslušnými vlastními vektory. První hl. komponenta je dána vztahem [7]:

$$y_1 = \sum_{j=1}^m v_{1j} x_j = v_1^T x \quad (7)$$

kde x jsou vstupní proměnné, v_1^T je vektor koeficientů.

Pro vektor koeficientů platí, že rozptyl pro první hlavní komponentu $D(y_1) = v_1^T$ je maximální. Vektory koeficientů jsou seřazeny od vektoru, co udává největší rozptyl hlavní komponenty $D(y_1)$ k vektoru co udává nejmenší rozptyl hlavní komponenty, tj. $D(y_1), D(y_2), \dots, D(y_m)$, kde $D(y_1) > D(y_2) > \dots > D(y_m)$. Další komponenty jsou sestaveny stejným postupem.

Schéma výpočtu hlavních komponent je na obrázku 2.1.



Obr. 2.1: Schéma výpočtu hlavních komponent

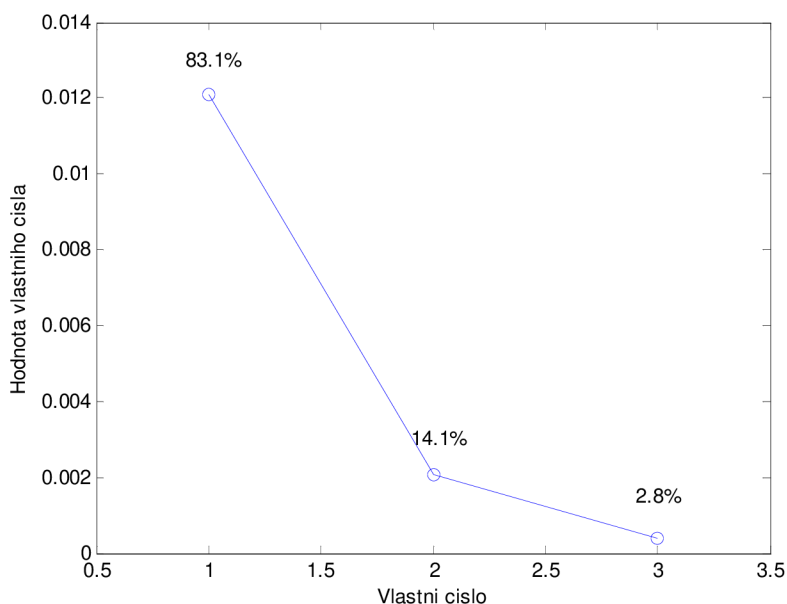
Jak již bylo zmíněno, první hlavní komponenta udává největší část rozptylu ze vstupních dat. Druhá hlavní komponenta udává největší část rozptylu, který není obsažen v první hlavní komponentě. Třetí hlavní komponenta udává největší část rozptylu, který není obsažen v první ani druhé hlavní komponentě, atd.

Z výše uvedených informací tedy vyplývá, že každá komponenta obsáhne určitou informaci ze vstupních dat, vyčerpá určitou variabilitu. Tato hodnota se obvykle uvádí v procentech. Při znalosti těchto údajů o vyčerpané variabilitě lze zvolit pouze několik hlavních komponent a dále pracovat jen s nimi. Obvykle se považuje za dostačující, když je vyčerpáno 90% variability. Tato hodnota se ale samozřejmě může měnit podle

účelu potřeby a využití PCA. Využívá se toho i při zpracování signálů, kdy neužitečné složky signálu (typicky šum) jsou obsaženy v komponentách, které vyčerpají zbývajících 10% variability.

K rozdělení komponent na užitečné a neužitečné, resp. výběru komponent, které vyčerpají požadovanou variabilitu, slouží tzv. scree plot.

Scree plot je graf, který udává velikost jednotlivých vlastních čísel a procentuální hodnotu vyčerpané variability jednotlivých komponent. Z obrázku 2.1 vyplývá, že první komponenta vyčerpá 83,1%, druhá 14,1% variability a třetí 2,8%. Je otázkou, kolik v tomto případě zvolit hlavních komponent. Neexistuje přesná hodnota, která by udávala, kolik procent variability je nezbytně nutných. Záleží to na experimentálních datech. Konkrétně v tomto případě je jisté, že se bude volit jedna nebo dvě komponenty. Kromě samotné procentuální hodnoty existuje ještě jedno kritérium, které může být použito, pokud nestačí vyčerpaná variabilita. Je to tvar křivky scree plotu. Dokud křivka klesá strmě dolů, můžou být komponenty vybírány dále. Pokud se začne narovnávat, je vhodné výběr ukončit [10]. Na obrázku 2.2 mezi první a druhou komponentou křivka strmě klesá, mezi druhou a třetí už se narovnáává, proto pokud v tomto případě budou zvoleny první dvě komponenty, tak výsledek rozhodně nebude chybný. Ale jak bylo řečeno, může se stát, že by byla dostačující pouze jedna komponenta. Vliv vyčerpané variability na výsledky klasifikace bude testován v praktické části.



Obr. 2.2: Ukázka-scree plot

2.2 Využití PCA při zpracování EKG

Při zpracování signálu se řeší různé problémy, jako je komprese dat, redukce šumu, určení důležitých příznaků atd. Důvodem je velký objem dat, obsahující informaci, která není důležitá pro interpretaci a klasifikaci. K řešení těchto problémů se dá použít PCA.

Když je signál rekurentní (jako např. EKG), analýza je často založena na vzorcích z opakujícího se úseku v různých periodách. U EKG je typickým příkladem srdeční cyklus. PCA při zpracování EKG signálu bere jako počáteční bod některý ze vzorků úseku, který se nachází v některé vhodné části srdečního rytmu. Umístění bodu v rámci srdečního rytmu může být u každé prováděné analýzy jiné a může zahrnovat celý srdeční rytmus nebo konkrétní aktivity, jako je P vlna, QRS komplex, nebo T vlna[6].

Je tedy třeba zvolit vhodná reprezentativní data (příznaky), která budou popisovat sledovaný signál. PCA může být použita ke zvolení takových vhodných příznaků. V tabulce 2.1 jsou uvedeni někteří vybraní autoři, kteří použili k výpočtu příznaků PCA. Také je uvedena úspěšnost klasifikace s využitím různých klasifikátorů.

Tab. 2.1: Přehled úspěšnosti klasifikace při použití PCA k výpočtu příznaků

AUTOŘI, ROK	VÝPOČET PŘÍZNAKŮ	KLASIFI-KÁTOR	ZAMĚŘENÍ	ÚSPĚŠNOST KLASIFIKACE
Ceylan R., Ozbay Y., 2007 [7]	PCA	NN	diagnostika srdeční poruchy	99,1%
Jankowski S., Dusza J. J., Wierzbowski M., Oreziak A., 2004 [11]	PCA	NN	srdeční arytmie	98,8%
Imah M. E., Afif F. A., Fanany M. I., Jatmiko W., Basaruddin T., 2011 [12]	PCA	SVM	prevence infarktu	94,5%
Hamoud M., Moussaoui A., 2010 [13]	PCA	SVM	rozdělení srdečních arytmií na 5 typů	98,7%
Murugan S., Radhakrishnan S., 2010 [14]	PCA	SVM	rozpoznávání ischemie	86,7%

Pro porovnání tabulka 2.2 obsahuje také jiné metody použité pro získání příznaků a úspěšnost klasifikace s jejich využitím.

Tab. 2.2: Přehled používaných metod

AUTOŘI, ROK	VÝPOČET PŘÍZNAKŮ	KLASIFI-KÁTOR	ÚSPĚŠNOST KLASIFIKACE
Palreddy S., Hu H. Y., Mani V., Tompkins W. J., 1997 [15]	LVQ	SOM	98,6%
Linh T. H., Osowski S., Stodolski M., 2003 [16]	WT	NN	96%
Parsad G. K., Sahambi J. S., 2003 [17]	WT	NN	96,8%
Chazal P., O'dwyer M., Reilly R. B., 2004 [18]	Morfologické	LDA	97%
Rodríguez J., Goni A., Illarramendi A., 2005 [19]	Morfologické	DT	96,1%

Příznaky:

PCA = principal component analysis, tj. analýza hlavních komponent

LVQ = learning vector quantization, tj. vektorová kvantizace učení

WT = wavelet transform, tj. vlnková transformace

Klasifikátory:

ANN = artificial neural network, tj. umělá neuronová síť

SVM = support vector machine, tj. metoda podpůrných vektorů

SOM = self organizing maps, tj. samo organizující se mapy

LDA = linear discriminant analysis, tj. lineární diskriminační analýza

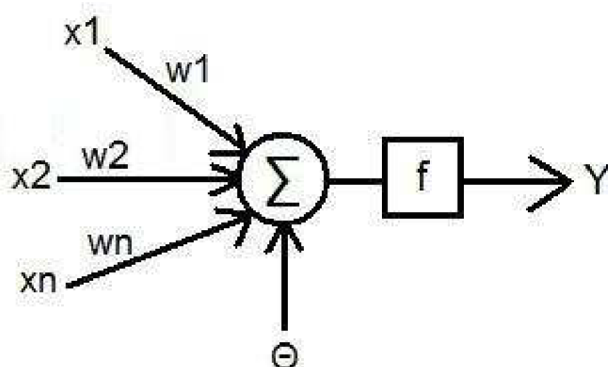
DT = decision tree, tj. rozhodovací stromy

3 UMĚLÉ NEURONOVÉ SÍTĚ

Umělé neuronové sítě (UNS) mohou být také označovány zkratkou ANN (z angl. artificial neural networks). Jejich prvotním úkolem bylo zjednodušeně modelovat základní procesy probíhající v mozku, typickými příklady jsou učení a paměť. Postupným vývojem UNS (větší složitostí a propracovaností) se modelování zdokonalilo. Proto mají UNS velké možnosti použití, mezi které patří i klasifikace dat, pro kterou budou použity v této práci [23].

3.1 Perceptron

Perceptron je umělý neuron a je základní stavební jednotkou neuronové sítě. Jeho nevýhodou je fakt, že lze aplikovat pouze na nejjednodušší problémy. To znamená pouze na data, která jdou lineárně oddělit. Jednoduchý perceptron (obr. 3.1) může mít n vstupů (x_1, x_2, \dots, x_n), které jsou váhovány odpovídajícími vahami (w_1, w_2, \dots, w_n). Například váha w_1 udává, jak významný je vstup x_1 , podobně pro ostatní váhy.



Obr. 3.1: Jednoduchý perceptron

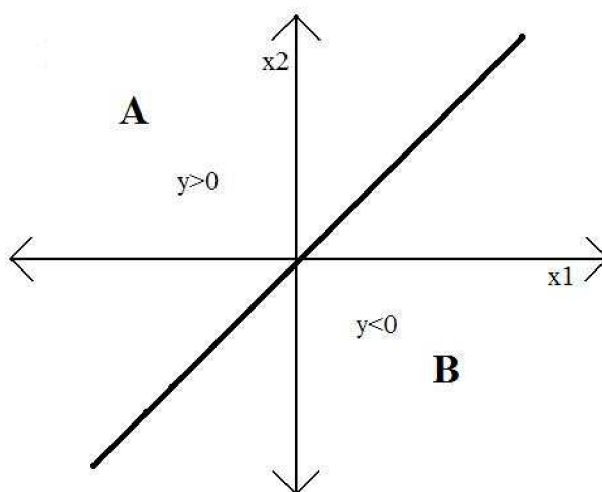
Výstupem je hodnota Y , která je dána váženým součtem vstupních hodnot, je určena vztahem [24]:

$$Y = f\left(\sum_{i=1}^k x_i * w_i - \Theta\right) \quad (8)$$

kde f je aktivační funkce neuronu, x_i je i -tý vstup, w_i je odpovídající i -tá váha, Θ je práh aktivační funkce neuronu a Y je výstup.

Využití neuronu jak klasifikátoru

Pro vysvětlení klasifikace pomocí neuronu je nejjednodušší zvolit neuron se dvěma vstupy a jedním binárním výstupem (vstupy budou zařazeny do jedné ze dvou tříd). Parametry neuronu jsou tedy vstupy x_1 a x_2 , jejich odpovídající váhy w_1 , w_2 , práh aktivační funkce a binární výstup y , který může nabývat hodnot 1 (odpovídající třídě A) nebo -1 (odpovídající třídě B). Tento jednoduchý neuron může sloužit ke klasifikování dvojic vstupů do dvou tříd (obr. 3.2). Sklon a pozice dělicí přímky, která od sebe odděluje třídy, je dána nastavením vah a aktivační funkce neuronu [24].



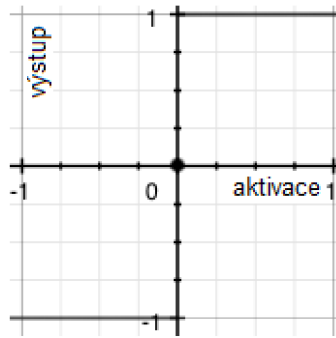
Obr. 3.2: Rozdělení do dvou tříd s využitím jednoho neuronu

Aktivační funkce

Aktivační funkce je charakteristika neuronu, která udává závislost výstupu neuronu na aktivaci. Ve většině případů je aktivační funkce rostoucí, mezi nejpoužívanější typy patří funkce skoková a sigmoidní [24].

Skoková funkce

V tomto případě je zvolen práh, a pokud je aktivační hodnota menší než práh, na výstupu bude hodnota -1. Při hodnotě větší než práh bude výstupní hodnota rovna +1. Ukázka skokové aktivační funkce je na obrázku 3.3.



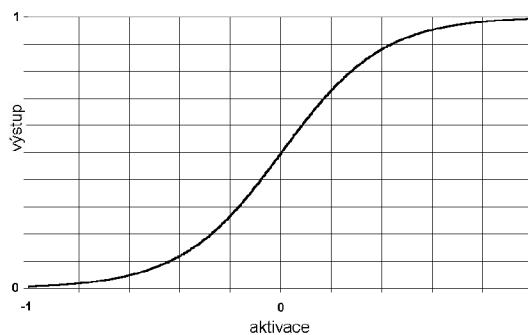
Obr. 3.3: Skoková aktivační funkce

Sigmoidní funkce

Při použití sigmoidní funkce je výstupní hodnota vypočítána podle vztahu [23]:

$$f(x) = \frac{1}{1 + e^{-ax}} \quad (9)$$

kde a je zvolený parametr, udávající tvar křivky. Funkce je vždy spojitá a má spojitě derivace (obr. 3.4).



Obr. 3.4: Sigmoidní funkce

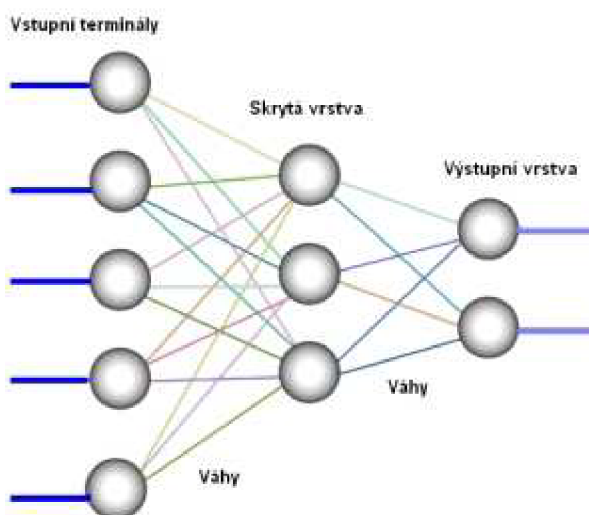
Typ aktivační funkce se volí podle toho, jaký typ dat je požadován na výstupu sítě. Pokud jsou požadovány na výstupu pouze dvě hodnoty (např. 0 a 1), je vhodné použít skokovou funkci. Když je požadován spojitý výstup v rozmezí $(-\infty; \infty)$, používá se lineární funkce. Sigmoidní funkce se používá při požadování spojitěho výstupu, když jsou vstupní hodnoty normalizovány v rozmezí $\langle -1; 1 \rangle$, což obvykle bývá ve skryté vrstvě vícevrstvé neuronové sítě.

3.2 Topologie sítí

Spojováním a vrstvením neuronů je možné vytvořit velké neuronové sítě. Podle toho, jak budou neurony pospojovány a vrstveny mohou vzniknout různé typy sítí. Topologie sítě se obvykle volí podle potřeby, tedy s ohledem na konkrétní experiment a data. Používají se dva základní typy: dopředné a rekurentní [24].

Dopředné sítě

Dopředné sítě (neboli feed-forward sítě) nevyužívají zpětnou vazbu mezi neurony. Jednou z nejpoužívanějších dopředných sítí je *vícevrstvá neuronová síť* (obr. 3.5).



Obr. 3.5: Vícevrstvá neuronová síť s jednou skrytou vrstvou [25]

Vícevrstvá síť se skládá z jednoduchých neuronů, které musí být propojeny podle určitých pravidel, aby bylo dosaženo správné funkčnosti sítě. Jak napovídá název, síť se skládá z několika vrstev, konkrétně ze tří: vstupní, skryté a výstupní, přičemž skrytých vrstev může být i více. Počet skrytých vrstev a počet neuronů v jednotlivých skrytých vrstvách se stanovuje empiricky, tedy na základě určitých zkušeností. Počet neuronů ve vstupní vrstvě je roven délce příznakových vektorů (tj. počtu příznaků). Ve výstupní vrstvě je počet neuronů dán počtem klasifikačních tříd a způsobem kódování výstupní hodnoty. Pro binární kódování a dvě klasifikační třídy tedy stačí jeden neuron se skokovou funkcí, bude mít dvě výstupní hodnoty (třída 0 nebo 1). S použitím dvou neuronů s binárním kódováním mohou být rozlišeny čtyři klasifikační třídy (00, 01, 10, 11) atd. Každý neuron příslušné vrstvy je propojen se všemi neurony v následující

vrstvě. Informace je ale přenášena pouze v dopředném směru, jelikož se jedná o dopřednou neuronovou síť. Pro trénování NS se využívá tzv. *učení s učitelem*, viz kapitola 3.3.

Rekurentní síť

Tyto sítě využívají zpětnou vazbu mezi všemi neurony. Oproti dopředným sítím jsou rekurentní sítě výrazně složitější jak při navrhování, tak i při procesu učení. Do této skupiny patří například Hopfieldova a Kohonenova síť. Využívají tzv. *učení bez učitele* (viz kapitola 3.3). Sítě tohoto typu nebudou použity v této práci.

3.3 Učení neuronových sítí

Neuronové sítě mohou být učeny dvěma způsoby, je to buď učení s učitelem, nebo učení bez učitele [23].

Učení s učitelem

S tímto typem pracují dopředné NS. V procesu učení jsou využívány vektory vstupních příznaků a k těmto vektorům požadované výstupy. Tato data tvoří trénovací množinu, která tedy obsahuje vektory, o kterých je známo, do které třídy patří. Trénovací množina je použita v procesu učení, kdy jsou hodnoty parametrů (vah a prahů) jednotlivých neuronů nastavovány tak, aby na výstupu byly získány požadované hodnoty. Jak bylo řečeno, mezi nejpoužívanější dopředné sítě patří síť typu *back propagation* (učící se na základě zpětného šíření chyby). Učení sítě u tohoto typu probíhá následovně[23]:

- 1) nejprve je síti předložen vektor vstupních dat, na který zareagují všechny neurony a je získána odezva sítě,
- 2) následuje výpočet chybového vektoru, což je rozdíl mezi požadovaným výstupem a odezvou,
- 3) dalším krokem je výpočet chyby sítě (někdy také označována jako energie nebo výkon sítě) ve skryté vrstvě pomocí chybové funkce dané vztahem[23]:

$$E = \sum_j \frac{1}{2} (y_j - t_j)^2 \quad (10)$$

kde j je počet neuronů, y_j je odezva j -tého neuronu a t_j je požadovaná odezva j -tého neuronu.

4) posledním krokem je úprava vah každého neuronu podle vztahu [20]:

$$\Delta w_i = -\eta \frac{\partial E}{\partial w_i} + \mu w_i \quad (11)$$

kde η je koeficient učení, E je chybová funkce, μ je koeficient změny vah v předchozím kroku a Δw_i je změna synaptické váhy v předchozím kroku.

Je nutné dosáhnout co nejmenší hodnoty chybové funkce. Jak je vidět ze vztahu (11), pro výpočet minima chybové funkce se používá tzv. metoda gradientního sestupu, kdy pro výpočet nadcházející hodnoty chyby dochází k posunu z původní pozice do následující ve směru záporného gradientu. Výhodou je spolehlivost.

Učení bez učitele

Tento typ učení využívají rekurentní NS. V tomto případě jsou síti předložena vstupní data, ale nejsou k nim žádná odpovídající data požadovaná na výstupu. Tento princip je založen na předpokladu, že síť si ve vstupních datech sama najde podstatné příznaky.

4 KLASIFIKACE EXPERIMENTÁLNÍCH DAT

V této kapitole jsou popsána experimentální data, použitá pro klasifikaci, jejich předzpracování pomocí PCA a samotná klasifikace pomocí navržených neuronových sítí. Výsledky klasifikace jsou následně vyhodnoceny.

4.1 Popis experimentálních dat

Data byla získána z databáze dostupné na ÚBMI. Jedná se o signály získané v průběhu experimentu na Novozélandských bílých králících. Králíci byli uvedeni do celkové anestezie. Dále bylo nutné otevřít hrudník, aby mohlo být srdce vyjmuto a následně umístěno do nádobky s Krebs-Henseleitovým roztokem. První fáze pokusu trvala třicet minut a sloužila ke stabilizaci (st), z toho prvních patnáct minut bylo využito jako kontrolní část. Po skončení první fáze byla zastavena perfuze, což způsobilo globální ischemii. Po dalších patnácti minutách došlo znovu k obnovení perfuze. Celkem třikrát se opakovalo zastavení a spuštění perfuze. Bylo tedy celkově sedm fází: kontrolní, třikrát ischemická (i1, i2, i3) a třikrát reperfuzní (r1, r2, r3). Po dobu celého pokusu byly ze srdce snímány EKG signály pomocí třívodového ortogonálního systému se vzorkovací frekvencí 2000 Hz. Jednalo se o bezdotykovou metodu, tedy elektrody byly umístěny na stěnách lázně, v níž bylo srdce. Podrobnější popis experimentu je dostupný v [22].

4.2 Předzpracování dat

Z naměřených dat byly vybrány úseky bez artefaktů a následně rozděleny na úseky podle fáze experimentu. Dále bylo odstraněno kolísání nulové izolinie s využitím Lynnova filtru typu horní propust a mezní frekvencí 0,5 Hz pomocí softwaru dostupného na ÚBMI [8]. Po odstranění kolísání izolinie došlo k detekci R-vln pomocí detektoru dostupného na ÚBMI, který využíval vlnkové transformace [9]. Na závěr byly vybrány pouze úseky QRST, tedy došlo k odstranění vlny P, protože ta není pro tento pokus (klasifikace normálních a ischemických cyklů) podstatná. Úseku QRST bylo dosaženo tak, že se zjistila pozice R-vlny (jejího maxima) a začátek byl nastaven o 59 vzorků zpět a konec 450 vzorků vpřed. Tím byly získány úseky o délce 510 vzorků (tabulka 4.1).

Tabulka 4.1: Ukázka souboru dat s QRST segmenty (výběr dat z jednoho experimentu)

Fáze	svod X	svod Y	svod Z
'st'	<510x4373 double>	<510x4373 double>	<510x4373 double>
'i1'	<510x675 double>	<510x675 double>	<510x675 double>
'r1'	<510x940 double>	<510x940 double>	<510x940 double>
'i2'	<510x720 double>	<510x720 double>	<510x720 double>
'r2'	<510x908 double>	<510x908 double>	<510x908 double>
'i3'	<510x746 double>	<510x746 double>	<510x746 double>
'r3'	<510x987 double>	<510x987 double>	<510x987 double>

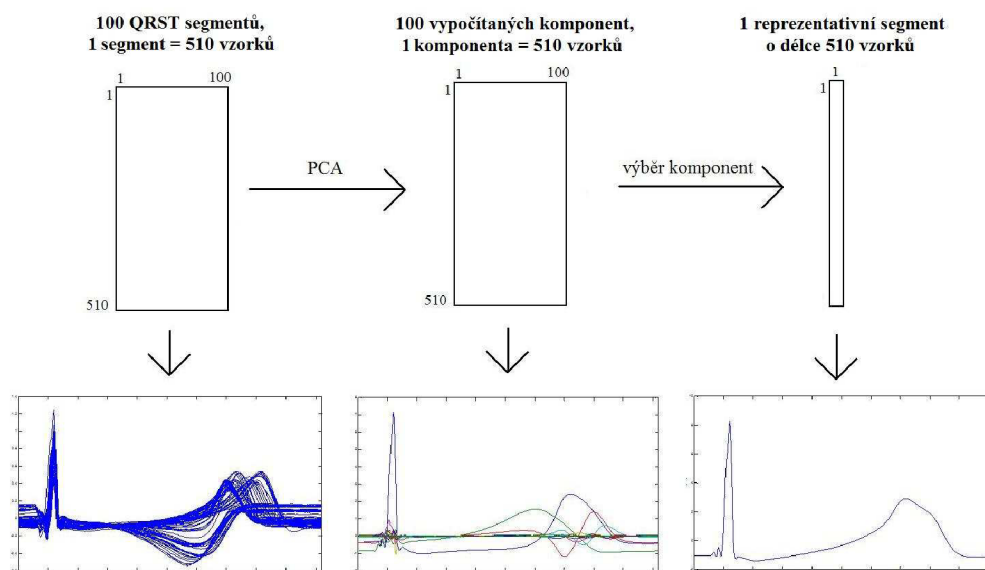
4.3 Zpracování experimentálních dat pomocí PCA

Jak je vidět v tabulce 4.1, všechna data jsou uložena v buňkách. Pro jednodušší práci s daty byla provedena úprava: všechny cykly z kontrolní fáze byly sloučeny do tří matic (pro každý svod zvlášť). Proběhlo sloučení dat ze dvou experimentů, což vedlo k získání celkem 9.000 QRST segmentů, tj. celkem 3 x 9.000 průběhů pro tři svody, každý průběh o délce 510 vzorků. To stejné bylo provedeno pro ischemickou fázi. V tomto případě se jedná asi o 7.000 srdečních cyklů popsaných třemi svody, celkem tedy 3 x 7.000 průběhů. Vzniklo tedy celkem šest matic QRST segmentů: *normálníX*, *normálníY*, *normálníZ* (každá matice o 9.000 průbězích), *ischemieX*, *ischemieY*, *ischemieZ* (každá matice o 7.000 průbězích). Z těchto matic budou vybírána data pro zpracování a klasifikaci.

Níže jsou uvedeny tři různé zvolené metody zpracování dat s využitím metody PCA (podrobně jsou popsány v kapitolách 4.3.1, 4.3.2 a 4.3.3), pomocí kterých je zkoumáno různé využití PCA. Liší se tím, že vstupem do PCA jsou vždy jiná data, resp. různé kombinace experimentálních dat, která byla popsána v předchozí kapitole. Jedná se konkrétně o následující využití:

1) Výběr reprezentativního segmentu ze 100 QRST segmentů

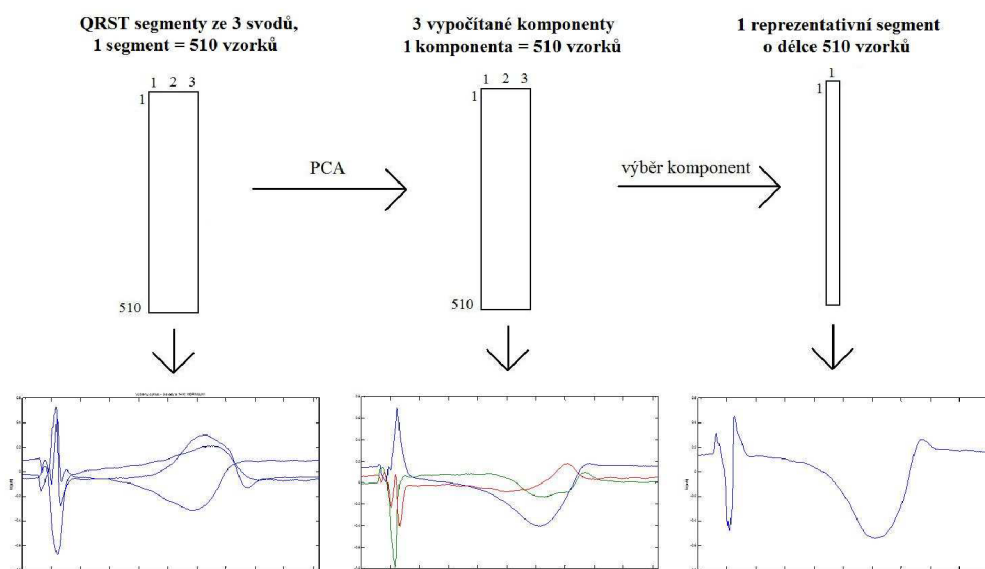
Ze vstupní matice dat, která obsahuje 100 segmentů o délce 510 vzorků, jsou vypočítány komponenty. Výběrem vhodného počtu komponent bude získán jeden reprezentativní segment o délce 510 vzorků (obr. 4.1).



Obr. 4.1: Schéma zpracování dat pomocí PCA: výběr 1 reprezentativního segmentu ze 100

2) Výběr reprezentativního segmentu pro popis QRST segmentů získaných ze tří svodů

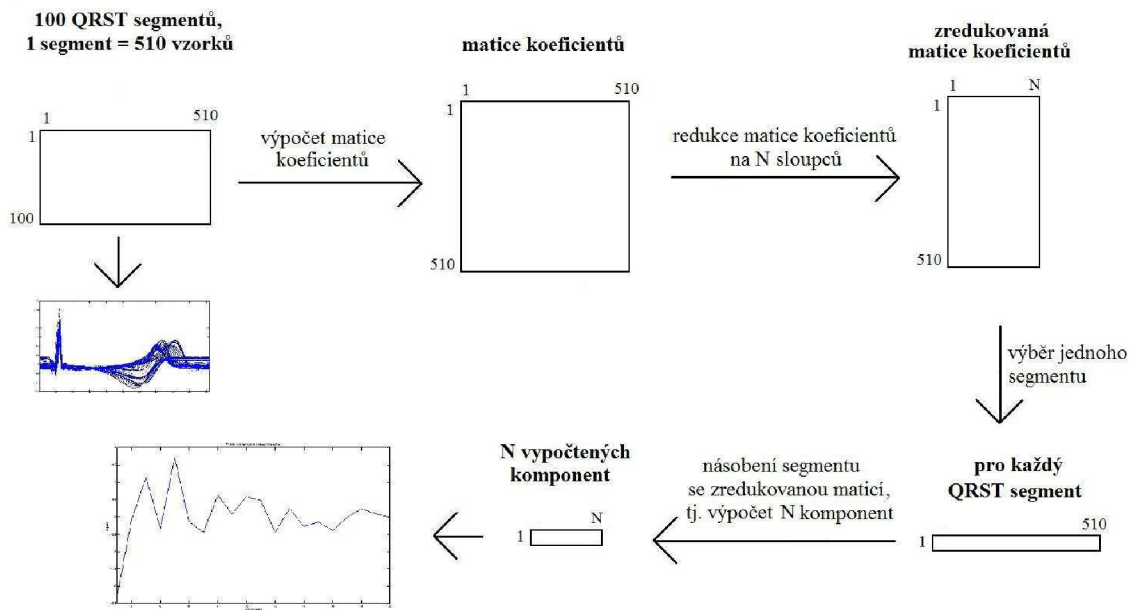
Základní myšlenka je stejná jako v předchozím případě, ale vstupní matice neobsahuje 100 segmentů získaných z jednoho svodu, ale 3 segmenty získané ze tří svodů (obr. 4.2).



Obr. 4.2: Schéma zpracování dat pomocí PCA: výběr 1 reprezentativního segmentu ze 3svodů

3) Redukce dat v rámci jednoho QRST segmentu

V tomto případě je matice koeficientů zredukována na požadovaný počet (N) sloupců. S využitím této zredukované matice je každý vstupní segment zredukován na zvolený počet (N) komponent (obr 4.3).



Obr. 4.3: Schéma zpracování dat pomocí PCA: redukce dat v rámci jednoho QRST segmentu

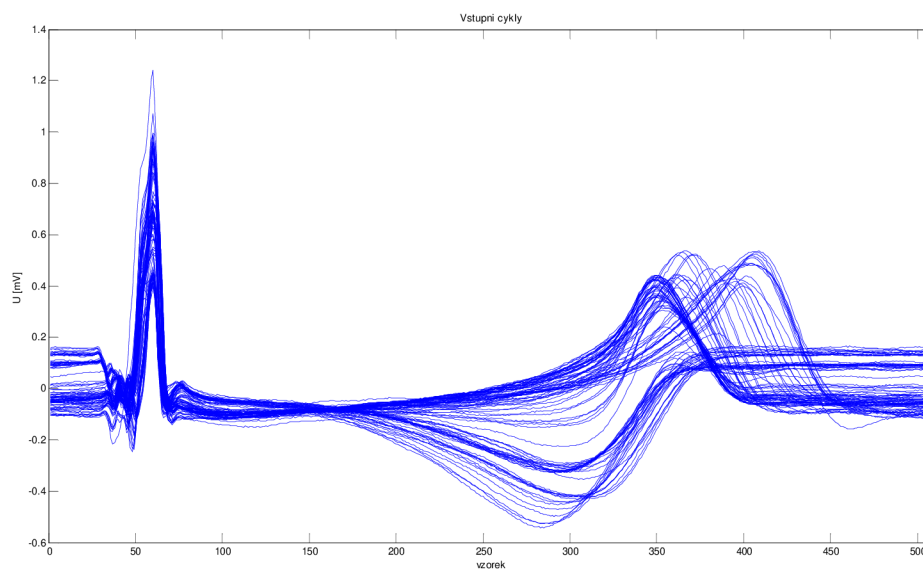
Ve všech třech případech jsou vstupní data podrobeny analýze hlavních komponent, kterou lze v programu Matlab provádět pomocí funkce *princomp*. Funkce má tři výstupy: koeficienty pro výpočet komponent (což je matice vlastních vektorů), vypočtené komponenty a také vlastní čísla. Z vlastních čísel je pomocí kumulované sumy zjištěno, kolik variability vyčerpá která hlavní komponenta. Procentuální hodnoty vyčerpané variability jsou zobrazeny ve scree plotech, na jejichž základě se rozhoduje o počtu komponent, se kterými se bude dále pracovat. Rozhodování o počtu je popsáno v kapitole 2.1.

Nyní už bude známo, která data jsou podstatná. Tato data mohou být zobrazena a uložena, aby mohla být využita pro další zpracování.

Níže jsou popsány výsledky zpracování experimentálních dat pomocí všech tří uvedených postupů.

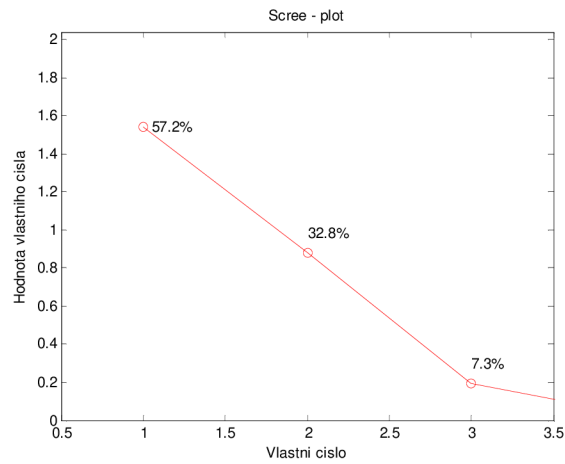
4.3.1 Aplikování PCA na sto cyklů z jednoho svodu

V tomto případě bude vstupem PCA sto srdečních cyklů. Požadavkem je získat jeden výsledný cyklus (komponentu), který obsáhne podstatnou část informace ze všech vstupních dat. Nejprve bylo zjištěno, kolik hlavních komponent bude vybráno. Ukázky pro tento případ (obr. 4.4, 4.5, 4.6) platí pro signál naměřený svodem X v kontrolní fázi experimentu. Vstupními daty do PCA je náhodně vybraných sto cyklů (obr. 4.4).



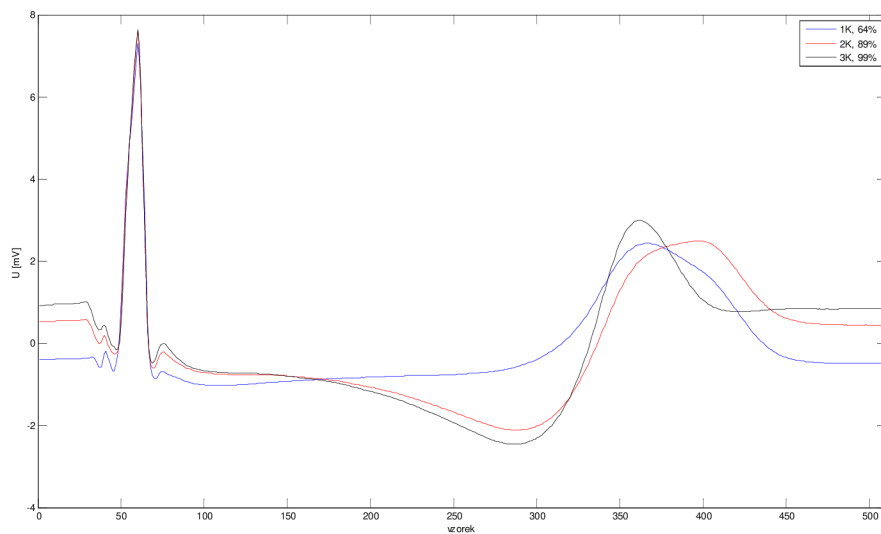
Obr. 4.4: Sto QRST segmentů vybraných z kontrolní fáze

Na dalším obrázku (4.5) je scree plot, ze kterého lze vyčíst, že první komponenta je nedostatečná, jelikož vyčerpá pouze 57% variability. Proto jsou zvoleny první dvě, které vyčerpají 90%. Zbývá informace (10%) bude označena jako nepodstatná a dále nebude brána v úvahu.



Obr. 4.5: Scree plot, sto cyklů

Na obrázku 4.6 lze pozorovat, jak vypadá výsledný signál, pro jednu, dvě nebo tři hlavní komponenty. Jelikož byly zvoleny dvě komponenty, výstup bude odpovídat průběhu zobrazeném červeně.



Obr. 4.6: Ukázka hlavních komponent (1k-57%, 2k-90%, 3k-97% variability)

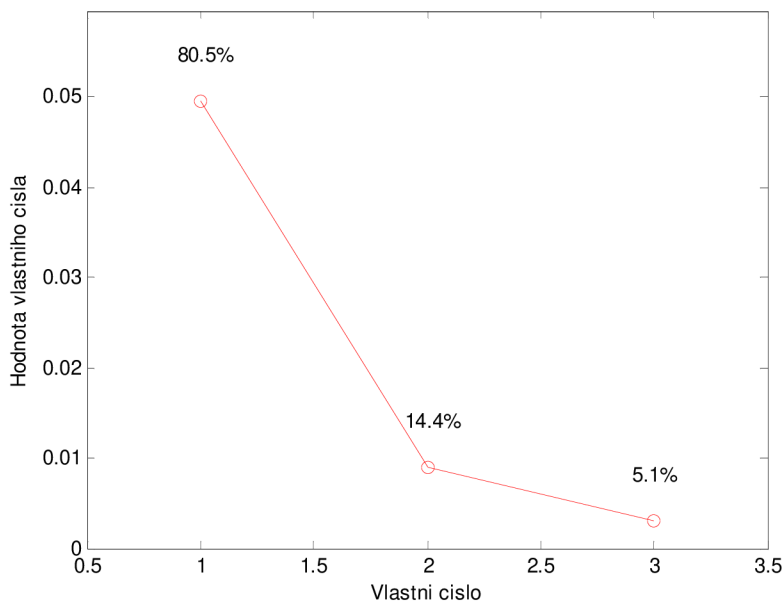
Opakováním výše uvedeného postupu byly získány celkem dvě množiny dat, přičemž každá obsahovala 50 průběhů pro každou fázi experimentu. Z jedné množiny dat byla vytvořena trénovací matice a z druhé testovací matice. Obě dvě mají 510 řádků (délka cyklů) a 100 sloupců (50 + 50 výstupů PCA). V lichých sloupcích budou data z normální fáze, v sudých sloupcích data z ischemické fáze. Ještě je nutné předdefinovat vektor, který bude mít tolik hodnot, kolik je vstupů – tedy 100. Na liché pozice bylo

zvoleno číslo 1, na sudé pozice 0, přičemž jednička znázorňuje normální a nula ischemická data. Tento vektor slouží jako vzor ke vstupním datům pro závěrečné vyhodnocení klasifikace.

4.3.2 Zpracování tří svodů pomocí PCA

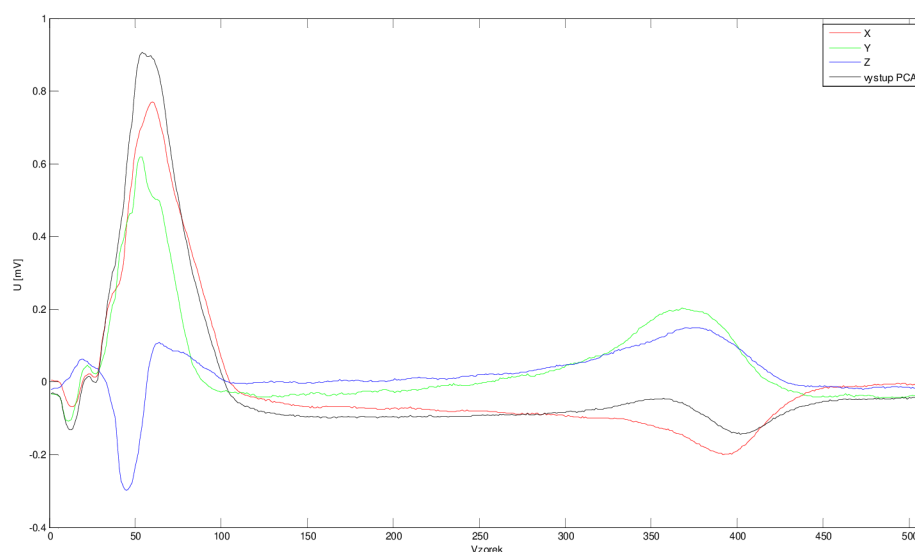
I v tomto případě byly rovněž vstupem data ze stejné fáze experimentu, ale byly zde využity všechny tři svody, kterými bylo srdce snímáno. V tomto případě byl výstup získán ze tří cyklů, nebo lépe řečeno ze tří průběhů, protože jeden srdeční cyklus byl monitorován pomocí tří svodů a byly tedy získány tři průběhy, které popisují jeden srdeční cyklus. Tyto tři průběhy jsou vstupem do PCA, výsledná hlavní komponenta tedy ponese informaci o jednom srdečním cyklu získanou ze tří signálů (svodů).

Zde byl pro ukázkou použit svod X ve fázi předpokládané ischemie. Opět byl vykreslen scree plot (obr. 4.7) pro určení vhodného počtu komponent. V tomto případě první komponenta vyčerpá 80,5% variability, což nemusí být dostatečné, proto budou zvoleny první dvě hlavní komponenty, které vyčerpají téměř 95%.



Obr. 4.7: Scree plot, tři svody

Na dalším obrázku (4.8) jsou vykresleny tři vstupní průběhy. Jedná se o cykly ze stejné fáze experimentu (předpokládané ischemie) zobrazené pomocí tří svodů: X (červeně), Y (zeleně) a Z (modře). Čtvrtá křivka znázorňuje výstup PCA (černě), který obsahuje 95% podstatné informace.



Obr. 4.8: Signály ze tří svodů (X, Y, Z) a výstup PCA

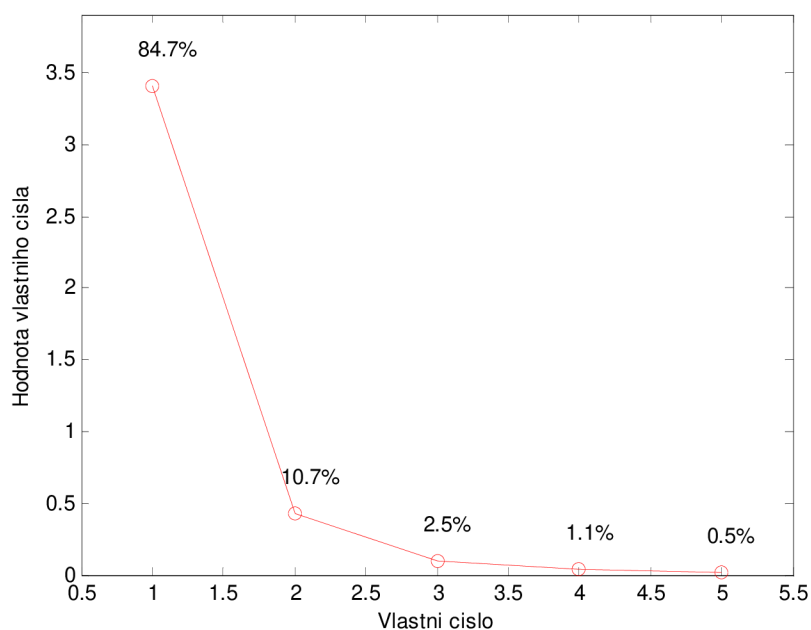
I v tomto případě při opakování popsaného postupu byly získány dvě množiny dat, ze kterých byla vytvořena testovací a trénovací matice. Každá obsahovala 50 průběhů z každé fáze. Také byl vytvořen vektor o sto hodnotách, který měl na lichých pozicích číslo 1 a na sudých číslo 0. Hodnoty reprezentují dvě klasifikační třídy (normální a ischemickou).

4.3.3 Využití PCA k redukci dat v rámci jednoho segmentu

V tomto případě nebyly využity přímo vypočtené komponenty pomocí funkce *princomp*, ale bude vytvořen postup pro jejich výpočet. V kapitole 4.3 je uvedeno, že funkce *princomp* má tři výstupy. V této části budou využity k výpočtu komponent vlastní vektory (neboli matice koeficientů), které jsou také výstupem zmíněné funkce. Hlavní komponenta je definována jako součin vstupních proměnných a odpovídajících koeficientů, jak je popsáno v kapitole 2. Koeficienty jsou uloženy v matici (vlastních vektorů), kde v prvním sloupci jsou hodnoty, ze kterých je počítána první komponenta, ve druhém sloupci koeficienty, ze kterých je počítána druhá komponenta a tak dále.

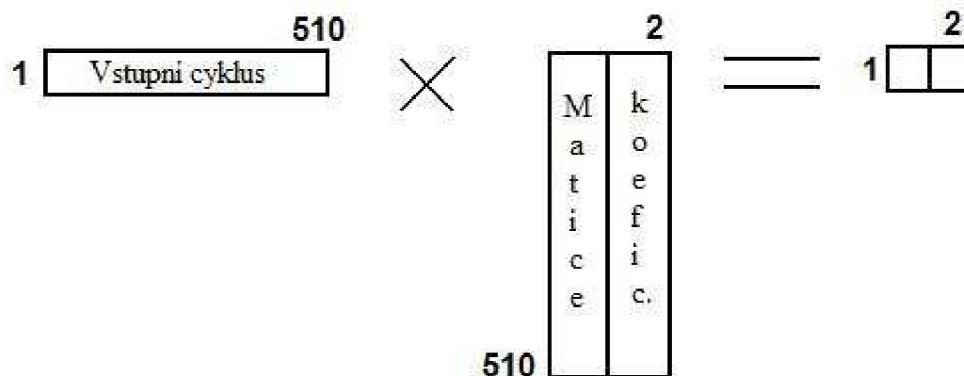
Při zvolení padesáti vstupních cyklů do PCA vznikne matice o velikosti 510 x 50 (jeden cyklus má 510 hodnot). Je však nutné matici transponovat kvůli tomu, aby bylo později

možné provést násobení s jinou maticí, což je popsáno níže. Tím tedy bude mít rozměry 50 x 510. Z těchto hodnot jsou vypočítány vlastní čísla a matice vlastních vektorů. Matice vlastních vektorů je čtvercová, má velikost 510 x 510, přičemž jednotlivé sloupce odpovídají vlastním vektorům. Nyní musí být zvolen požadovaný počet komponent, proto bude zobrazen scree plot (obr. 4.9), ze kterého lze usoudit, že podstatnou informaci nesou první dvě hlavní komponenty, vyčerpají 95,4%.



Obr. 4.9: Scree plot, redukce dat

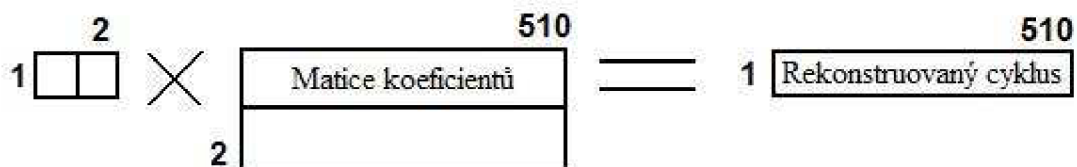
Zvoleny byly tedy dvě hlavní komponenty. Po určení požadovaného počtu komponent bude vybráno právě tolik vlastních vektorů. Matice koeficientů se tedy zredukuje, počet řádků bude pořád 510, ale počet sloupců bude roven zvolenému počtu komponent, v tomto případě tedy 2. Nyní dojde k násobení vstupních cyklů se zredukovanou maticí koeficientů. Vstupní cykly budou brány po jednom, násobit se tedy bude vektor 1 x 510 s maticí 510 x 2. Právě kvůli tomuto kroku bylo nutné matici vstupních cyklů transponovat, aby bylo splněno pravidlo pro násobení dvou matic nebo vektoru s maticí. To říká, že matice mohou být násobeny pouze v případě, že první má tolik sloupců kolik má druhá řádků. V tomto případě má vektor stejný počet sloupců jako má matice řádků (510), násobení je tedy možné (obr. 4.10).



Obr. 4.10: Násobení vstupního cyklu se zredukovanou maticí koeficientů

Došlo k vynásobení koeficientů pro první a druhou komponentu se vstupním cyklem. Výsledkem bude vektor 1 x 2, tedy pouze dvě hodnoty. To jsou dvě komponenty, pomocí kterých bude popsán celý cyklus. Takto se na vstupu vystřídají všechny vybrané cykly a každý bude popsán dvěma komponentami. Tím bylo dosaženo projekce vstupních dat do nových souřadnic, které jsou dány prvními dvěma komponentami, jedná se tedy o projekci vstupních vektorů pouze na dvě komponenty. Právě tímto došlo k redukci hodnot z 510 na 2.

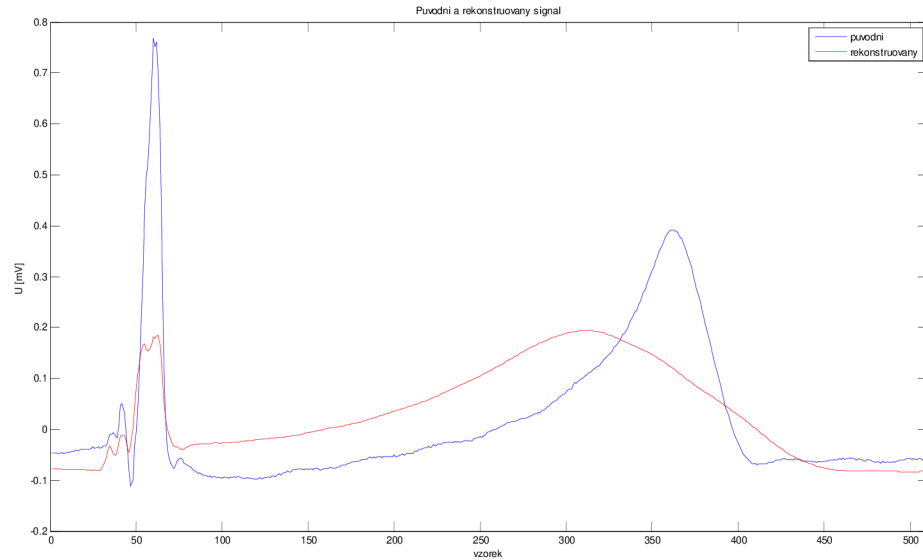
Nyní je však nutné ověřit, jestli nebyla data zredukována až příliš. Proto bude ze dvou hodnot, které nyní reprezentují cyklus zrekonstruován původní cyklus. Opět k tomu bude použita matice koeficientů, kterou budou vynásobeny dvě komponenty reprezentující cyklus. I u zpětné rekonstrukce je však nutné provést transponování matice, konkrétně matice koeficientů, aby bylo možné provést násobení. Násoben bude vektor 1 x 2 s maticí 2 x 510. Výsledkem tohoto násobení bude vektor 1 x 510 (obr. 4.11).



Obr. 4.11: Získání zrekonstruovaného cyklu pomocí zredukované matice koeficientů

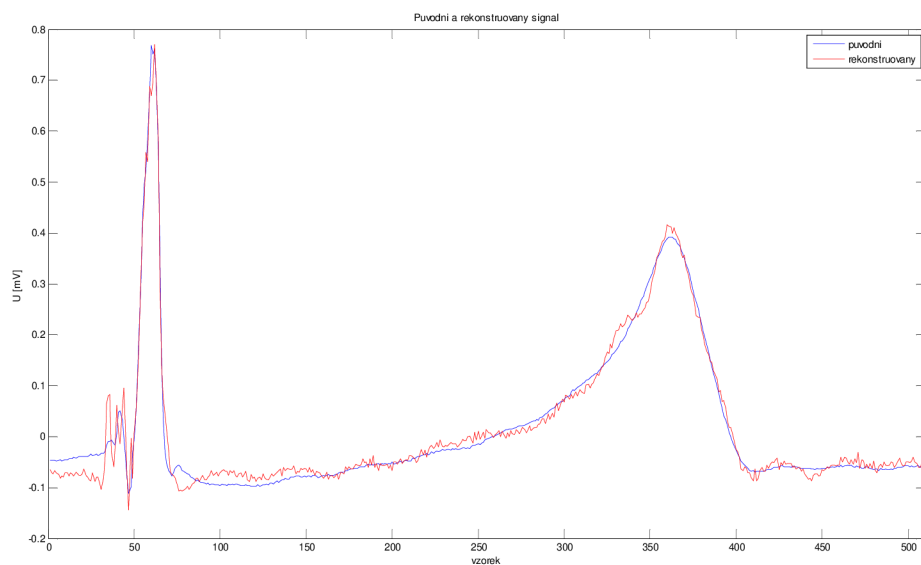
V této chvíli je už možné porovnat původní cyklus s zrekonstruovaným. V ideálním případě by měly být totožné nebo s minimálními rozdíly. Pro porovnání jsou oba cykly

zobrazeny v jednom grafu (obr. 4.12).



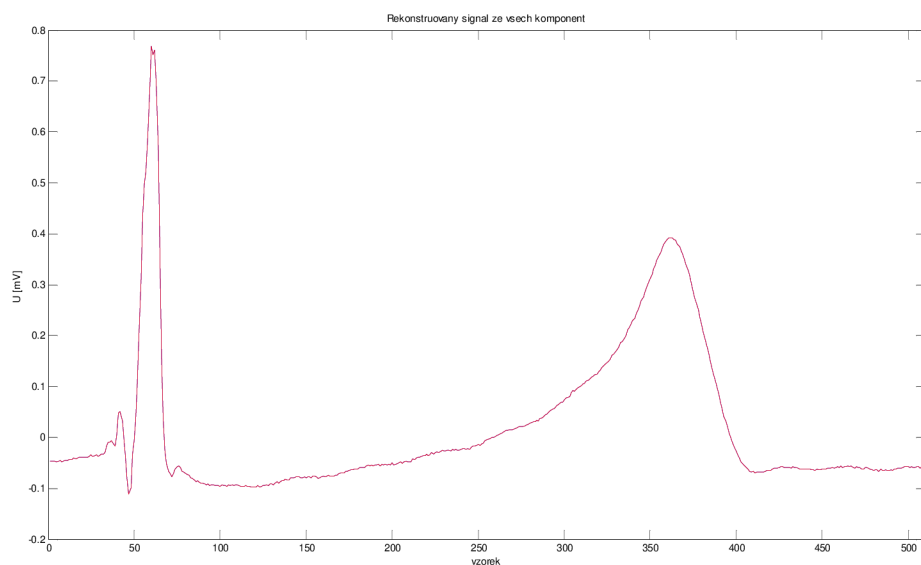
Obr. 4.12: Původní a rekonstruovaný cyklus (2 použité komponenty, korelace=0,73)

Z grafu je na první pohled zřejmé, že cykly se výrazně liší. To znamená, že data byla příliš zredukována, je tedy nutné výrazně zvýšit požadovaný počet komponent. Z tohoto příkladu lze také usoudit, že scree plot v tomto případě není vhodný pro určení vhodného počtu komponent. Je tedy třeba zvolit jiné kritérium pro určení minimálního počtu komponent. Nabízí se možnost využití korelace, přesněji výpočtu korelačního koeficientu mezi původním a rekonstruovaným cyklem. K výpočtu bude použita funkce *corrcoef*. Vypočtený korelační koeficient mezi průběhy v tomto případě je 73%. To je velmi málo a bude třeba tento koeficient zvýšit. Jak bylo řečeno, bude to provedeno větším počtem komponent. Je nutné volit kompromis, je třeba co nejméně komponent, ale s dostatečnou korelací mezi cykly. Na dalším obrázku (4.13) je původní a rekonstruovaný signál, ovšem v tomto případě byl požadovaný počet komponent nastaven na 20. Vizually jsou si signály velice podobné a korelace mezi nimi je 99%.



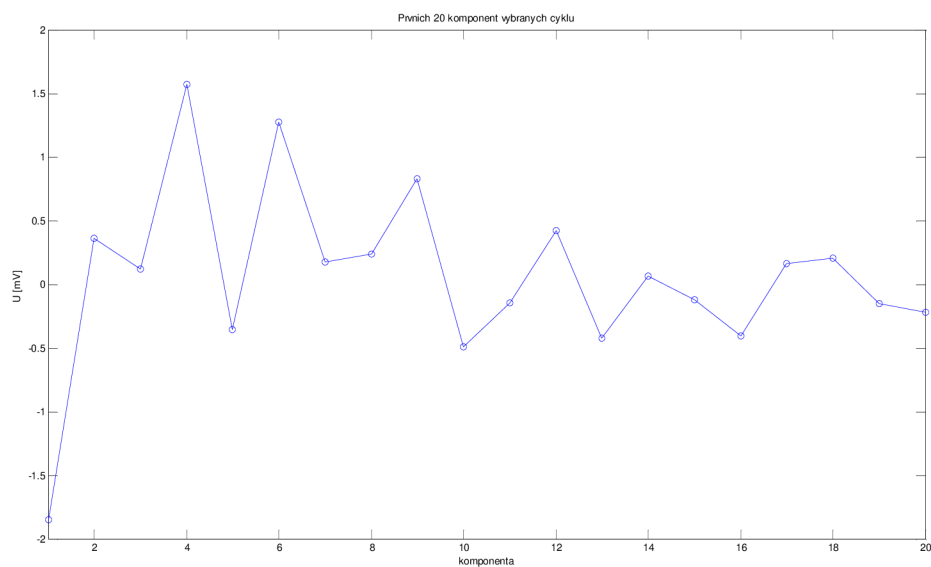
Obr. 4.13: Původní a rekonstruovaný cyklus (20 použitých komponent, korelace=0,99)

Tento výsledek už lze považovat za uspokojivý, protože signály jsou si podobné, korelační koeficient mezi nimi je roven 0,99 a došlo k redukci dat vstupního cyklu z původních 510 na 20. To, že průběh rekonstruovaného signálu působí „zašuměně“ je dáno mírnou ztrátou informace při redukci na 20 komponent. Pro ukázkou je na dalším obrázku (4.14) vykreslen cyklus rekonstruovaný ze všech komponent.



Obr. 4.14: Cyklus rekonstruovaný ze všech komponent (korelace = 1)

Cyklus tedy bude popsán pouze pomocí 20-ti hodnot, tedy prvních 20-ti komponent (obr 4.15).



Obr. 4.15: Prvních 20 komponent vstupního cyklu

Testovací a trénovací matice budou získány stejně jako v předchozích dvou případech, ale budou mít jiné rozměry. V každé matici bude 50 průběhů z každé fáze, takže počet sloupců zůstane stejný (100), ale počet řádků bude roven zvolenému počtu komponent. Požadovaný výstup bude opět reprezentován vektorem o sto hodnotách, ve kterém se budou střídát jedničky a nuly.

Komponenty získané ve výše uvedených případech tvoří vektory příznaků, je tedy možné přistoupit k návrhu klasifikátorů. Ukázky komponent pro normální a ischemické průběhy jsou uvedeny v kapitolách 4.5.1, 4.5.2 a 4.5.3.

4.4 Návrh klasifikátorů

V této kapitole je uvedeno, jaký vliv má změna počtu neuronů ve skryté vrstvě, počet skrytých vrstev a také počet epoch při učení NS na výslednou klasifikaci. Vždy se jedná o maximální možný počet epoch, který bude nastaven do NS, protože učení může skončit dříve (kapitola 4.5.1). Vstupní data do NS byla vždy stejná a s využitím *for-cyklu* byly měněny hodnoty parametrů a zaznamenávány výsledky klasifikace.

Použita byla vícevrstvá neuronová síť s jednou nebo dvěma skrytými vrstvami, využívající algoritmus back propagation pro natrénování sítě. Aktivační funkce mezi vstupní a skrytou vrstvou byla sigmoidní, mezi skrytou vrstvou a výstupní byla použita lineární aktivační funkce. V případě použití dvou skrytých vrstev byla mezi nimi použita sigmoidní aktivační funkce. NS byla naučena pomocí trénovací množiny dat, která obsahovala příznakové vektory (což byly vypočítané komponenty) a z hodnot požadovaných na výstupu. Pomocí naučené sítě byla testována testovací množina.

Výpočet přesnosti klasifikace se provádí pomocí vztahu[21]:

$$P = \frac{TP + TN}{TP + TN + FP + FN} * 100 \quad (12)$$

kde P je přesnost klasifikace, TP je počet skutečně pozitivních, TN počet skutečně negativních, FP počet falešně pozitivních a FN počet falešně negativních přiřazených případů do odpovídajících tříd.

Parametry pro aplikaci PCA na 100 cyklů

Níže uvedené tabulky uvádějí experimentálně získané hodnoty. Pro každý počet neuronů a počet epoch byly provedeny tři testy, každý test pro jiný svod. Jak je vidět v tabulce 4.2, úspěšnost klasifikace bylo vždy 100%, proto v tomto případě nebylo třeba testovat větší počet skrytých vrstev.

Tab. 4.2: Úspěšnost klasifikace při různém počtu neuronů 1

Počet neuronů	Úspěšnost klasifikace [%]			
	Test 1	Test 2	Test 3	Průměr
5	100	100	100	100
6	100	100	100	100
7	100	100	100	100
8	100	100	100	100
9	100	100	100	100
10	100	100	100	100
11	100	100	100	100
12	100	100	100	100
13	100	100	100	100
14	100	100	100	100
15	100	100	100	100

V tabulce 4.2 je možné vyčíst, že v tomto případě nezáleží na počtu neuronů. Klasifikace je ve všech případech stoprocentní. Proto v tomto případě mohl být počet neuronů libovolně zvolen, nastavena byla hodnota 10. V tabulce 4.3 se nachází údaje o úspěšnosti klasifikace při různém počtu epoch v procesu učení.

Tab. 4.3: Úspěšnost klasifikace při různém počtu epoch 1

Počet epoch		5	10	15	50	100	200
Úspěšnost klasifikace[%]	Test 1	92	97	100	100	100	100
	Test 2	83	91	99	100	100	100
	Test 3	91	92	100	100	100	100

Maximální počet epoch byl nastaven na 50 na základě tabulky 4.3. Jak je vidět, tak výsledek klasifikace dosáhne maxima nejpozději při padesáti epochách, proto není třeba volit vyšší číslo.

Parametry pro využití PCA na tři svody

V tabulkách jsou podobně jako v předchozím případě uvedeny získané hodnoty

provedením několika testů, každý pro jiné náhodně vybrané cykly. Z tabulky 4.4 vyplývá, že nejlepších výsledků bylo dosaženo, když byl počet neuronů nastaven na 9.

Tab. 4.4: Úspěšnost klasifikace při různém počtu neuronů 2

Počet neuronů	Úspěšnost klasifikace [%]			
	Test 1	Test 2	Test 3	Průměr
5	96	89	94	93,0
6	96	96	95	95,7
7	96	96	95	95,7
8	96	98	91	95,0
9	99	98	96	97,7
10	96	98	93	95,7
11	96	97	93	95,3
12	94	97	92	94,3
13	94	97	92	94,3
14	94	97	92	94,3
15	94	97	92	94,3

V tabulce 4.5 lze vyčíst, že pokud proběhne 300 epoch, tak výsledek klasifikace se už nezlepší. Proto bude jako maximální možný počet epoch nastaveno číslo 300.

Tab. 4.5: Úspěšnost klasifikace při různém počtu epoch 2

Počet epoch		50	100	200	300	400	500
Úspěšnost klasifikace[%]	Test 1	94	95	96	98	98	98
	Test 2	95	96	96	96	96	96
	Test 3	91	95	95	95	95	95

Při zvolení dvou skrytých vrstev místo jedné (každá obsahuje devět neuronů) nedochází k žádnému výraznému zlepšení, pouze o 0,3% (tab. 4.6). Proto bude počet skrytých vrstev nastaven na jednu.

Tab. 4.6: Úspěšnost klasifikace při různém počtu skrytých vrstev 1

		Úspěšnost klasifikace [%]			
		Test 1	Test 2	Test 3	Průměr
Počet skrytých vrstev	1	99	92	96	95,7%
	2	98	92	98	96%

Parametry pro redukci v rámci jednoho cyklu

I v tomto případě jsou experimentálně získané hodnoty uvedeny v tabulkách. Rovněž byly provedeny testy pro všechny parametry, pro každý svod jeden. Nejlepších výsledků klasifikace v tomto případě dosahuje NS, pokud je počet neuronů nastaven na 10 (tab. 4.7).

Tab. 4.7: Úspěšnost klasifikace při různém počtu neuronů 3

Počet neuronů	Úspěšnost klasifikace [%]			
	Test 1	Test 2	Test 3	Průměr
5	72	81	60	71,0
6	74	83	61	72,7
7	75	87	61	74,3
8	75	87	62	74,7
9	75	87	66	76,0
10	77	89	66	77,3
11	77	87	62	75,3
12	76	85	62	74,3
13	75	85	61	73,7
14	75	84	60	73,0
15	75	83	60	72,7

Maximální počet epoch byl v tomto případě stanoven na 300 na základě údajů v tabulce 4.8.

Tab. 4.8: Úspěšnost klasifikace při různém počtu epoch 3

Počet epoch		50	100	200	300	400	500
Úspěšnost klasifikace[%]	Test 1	52	59	60	64	65	65
	Test 2	60	61	66	71	71	71
	Test 3	67	69	70	75	75	75

Z tabulky 4.9 vyplývá, že ani v tomto případě navýšení počtu skrytých vrstev (v každé bylo deset neuronů) nepřineslo zlepšení úspěšnosti klasifikace, proto i v tomto případě bude použita jedna skrytá vrstva.

Tab. 4.9: Úspěšnost klasifikace při různém počtu skrytých vrstev 2

		Úspěšnost klasifikace [%]			
		Test 1	Test 2	Test 3	Průměr
Počet skrytých vrstev	1	67	53	98	72,7%
	2	69	52	96	72,3%

4.5 Aplikace pro zpracování dat pomocí PCA a klasifikaci pomocí UNS

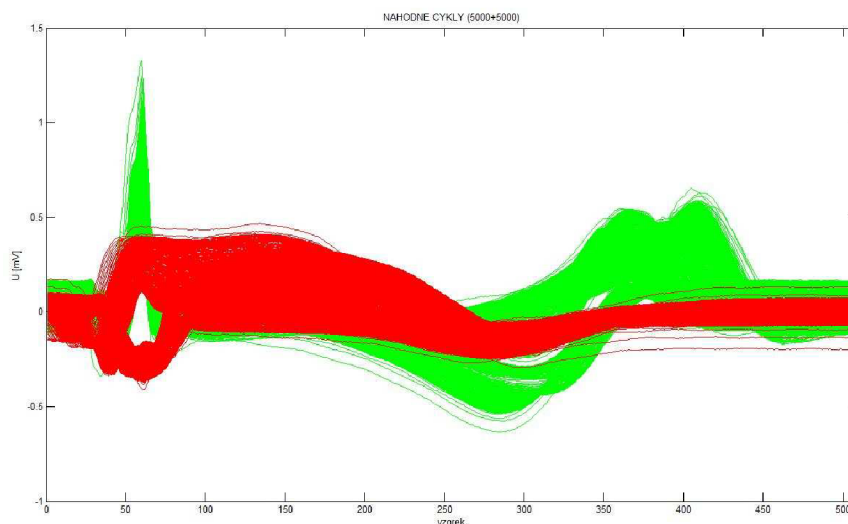
V kapitole 4.3 jsou uvedeny tři příklady použití metody hlavních komponent. Každá metoda je specifická, proto byly vytvořeny tři aplikace (pro každou metodu jiná) s grafickým uživatelským rozhraním. Všechny aplikace obsahují několik jednoduchých prvků (tlačítka, výběrová menu, grafy) a jsou navrženy tak, aby mohly být uživatelem snadno ovládány.

Aplikace se skládají ze dvou částí. První je programovaná část, která obsahuje skript se všemi příkazy, které budou využívány. Druhá část obsahuje grafické prvky. Ty udávají, kdy bude aktivována konkrétní část programu a také zobrazují výsledky. Ke spuštění aplikace jsou nutné obě části, jedna s příponou *.m* a druhá s příponou *.fig*. Tyto části musí být umístěny ve stejné složce jako vstupní data. Podrobněji jsou jednotlivé aplikace popsány níže.

4.5.1 Aplikace s použitím PCA na 100 cyklů

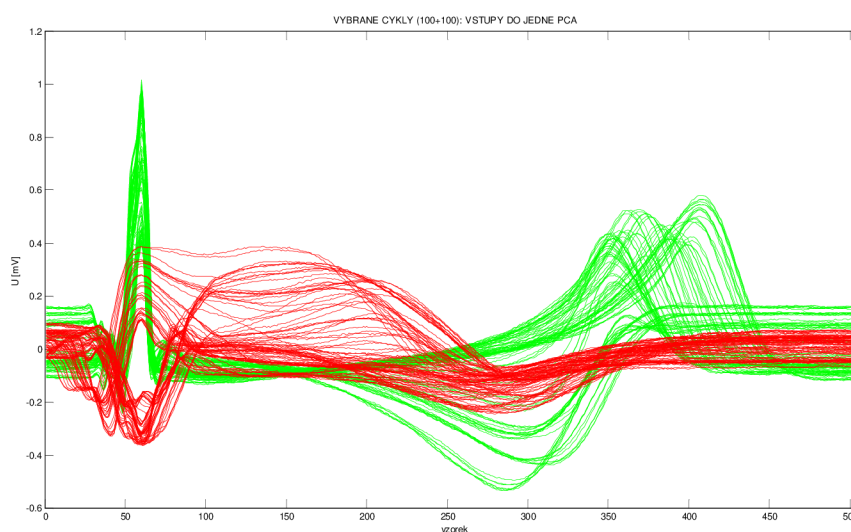
Tato aplikace se spouští otevřením souboru *jedenZEst.m* a jeho následným spuštěním tlačítkem *run*. Zobrazí se pouze tlačítko *Nacist data*. Data, která obsahují normální i ischemické průběhy budou načtena ze souborů a uložena do proměnných, se kterými bude dále pracováno. Až budou data načtena, dojde k odkrytí dvou částí. Uživatel si může zvolit jeden ze tří svodů (X, Y, Z), se kterým se bude dále pracovat. Pomocí tlačítka *Nahodne vybrat cykly* bude náhodně vybráno 5.000 normálních průběhů (z celkových 9-ti tisíc) a 5.000 ischemických průběhů (z celkových 7-mi tisíc), které budou vykresleny do prvního grafu (obr 4.16). Toto číslo je zvoleno s ohledem na pozdější klasifikaci dat. Do klasifikace bude vstupovat 50 normálních

a 50 ischemických průběhů, které budou výstupy PCA. Pro získání jednoho výstupu PCA (tj. komponenty) bude použito 100 naměřených cyklů, tedy pro získání padesáti výstupů PCA bude třeba 5.000 naměřených cyklů (50 x 100) z každé fáze experimentu.



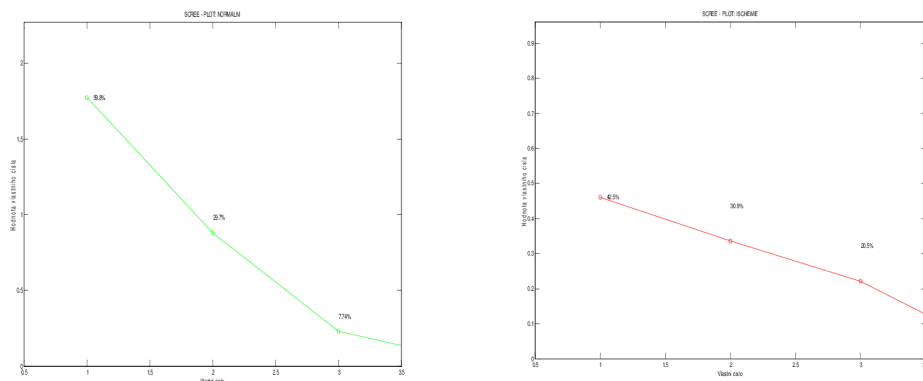
Obr. 4.16: Náhodně vybrané normální (z) a ischemické (č) cykly (5000 + 5000), svod X

Ve druhém grafu bude zobrazeno sto normálních cyklů, což je vstup do jedné PCA pro normální cykly. To stejné bude vykresleno pro ischemické cykly, tedy sto vstupních cyklů pro jeden průběh PCA pro ischemickou část (obr. 4.17).



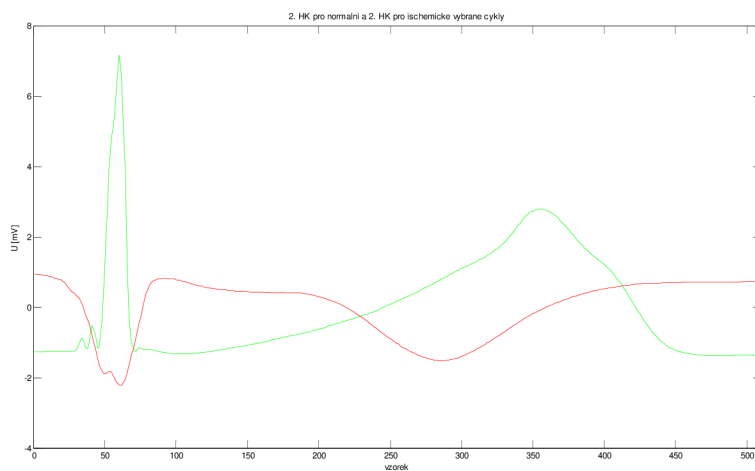
Obr. 4.17: Vybrané ukázkové cykly: normální a ischemické

V tomto okamžiku je možné použít tlačítko *Scree ploty*. Jak jeho název vypovídá, dojde k zobrazení dvou scree plotů, jeden pro vybrané normální cykly a druhý pro vybrané ischemické cykly (obr. 4.18). Na základě zobrazených scree plotů nyní bude uživatelem zvolen vhodný počet komponent, se kterými se bude dále pracovat. Je možno vybrat jednu, dvě nebo tři komponenty. Bylo vyzkoušeno, že tři komponenty vyčerpají nejméně 90% variability, proto není třeba volit více. Jak optimálně zvolit počet komponent je popsáno v kapitole 2.2.

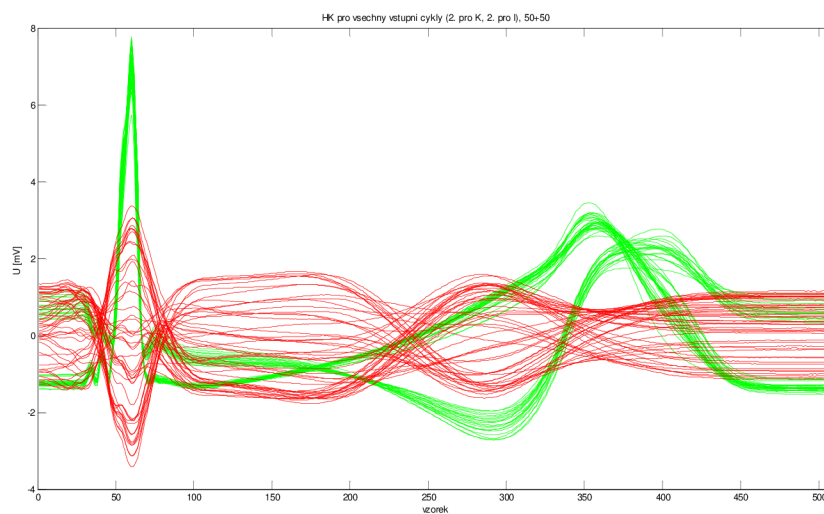


Obr. 4.18: Scree ploty pro normální (vlevo) a ischemická data (vpravo)

Stiskem tlačítka *Zobrazit vysledky PCA* bude provedena analýza hlavních komponent pomocí funkce *princomp*, která vypočítá komponenty. Po té budou v pátém grafu zobrazeny výstupy PCA pro ukázkové vybrané cykly (obr. 4.19). Jelikož PCA pro každou fázi proběhla padesátkrát, budou do posledního grafu vykresleny výsledky všech PCA, tedy 50 a 50. To jsou vstupy do neuronové sítě (obr. 4.20).

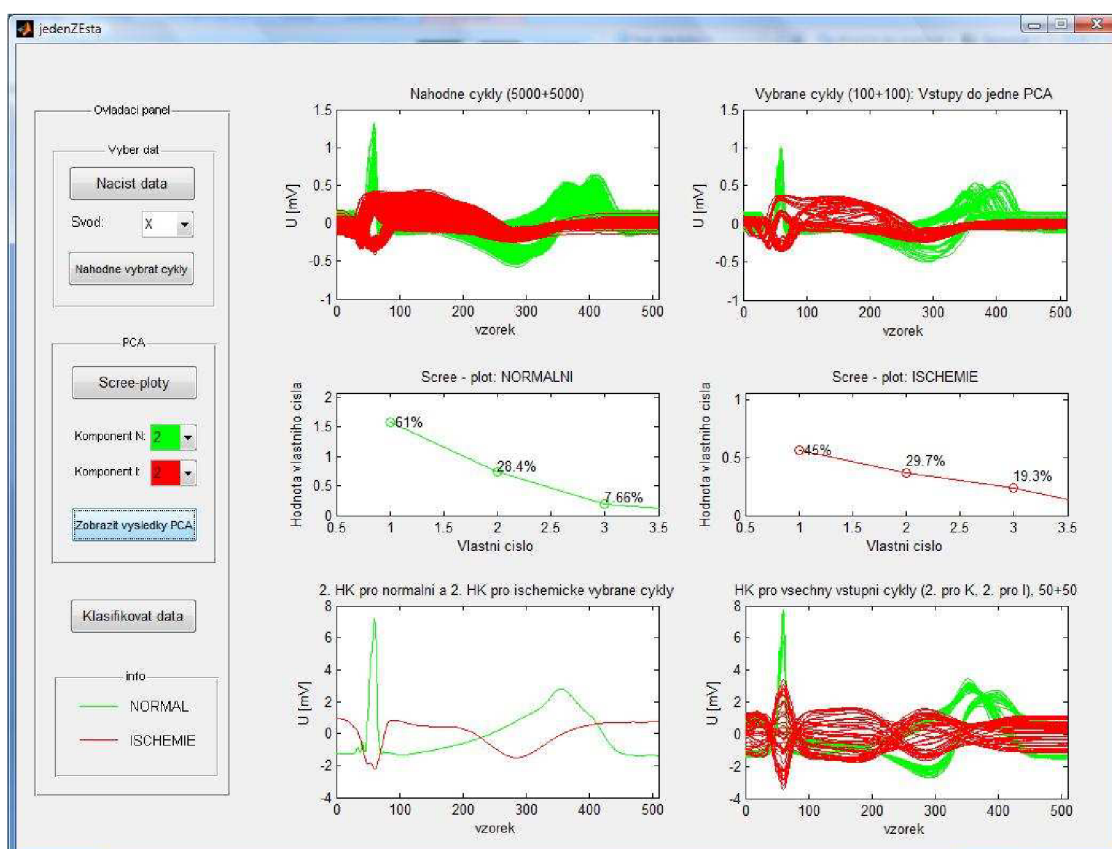


Obr. 4.19: První dvě hlavní komponenty pro vybrané cykly - normální (z) a ischemické (č)



Obr. 4.20: První dvě komponenty všech cyklů: normálních (z) a ischemických (č)

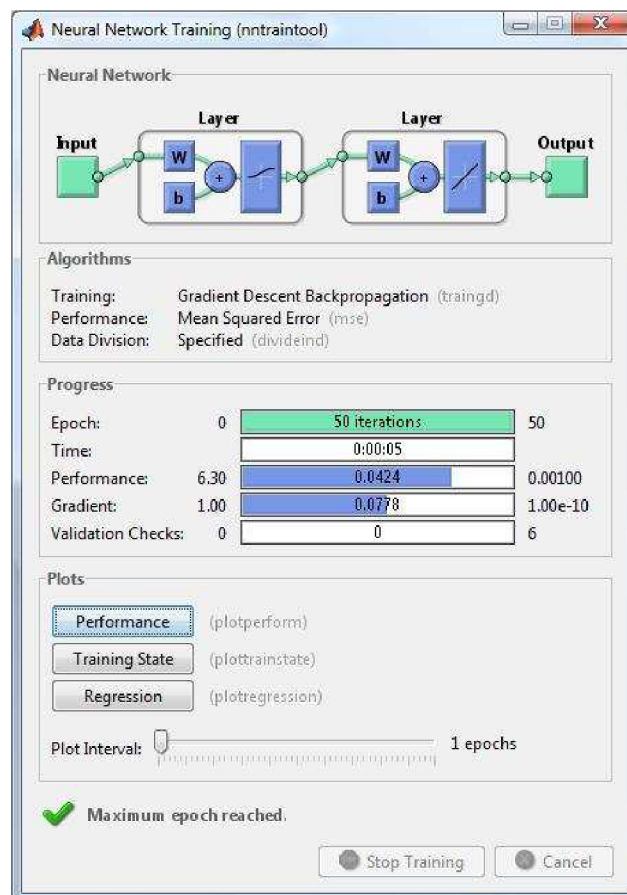
Jako bylo řečeno v kapitole 4.3.1, byly tedy vytvořeny dvě matice: jedna trénovací a druhá testovací, kvůli větší variabilitě dat. V aplikaci (obr. 4.21) jsou však zobrazována pouze data z jedné množiny (testovací) kvůli lepší přehlednosti.



Obr. 4.21: Aplikace Jeden ze sta

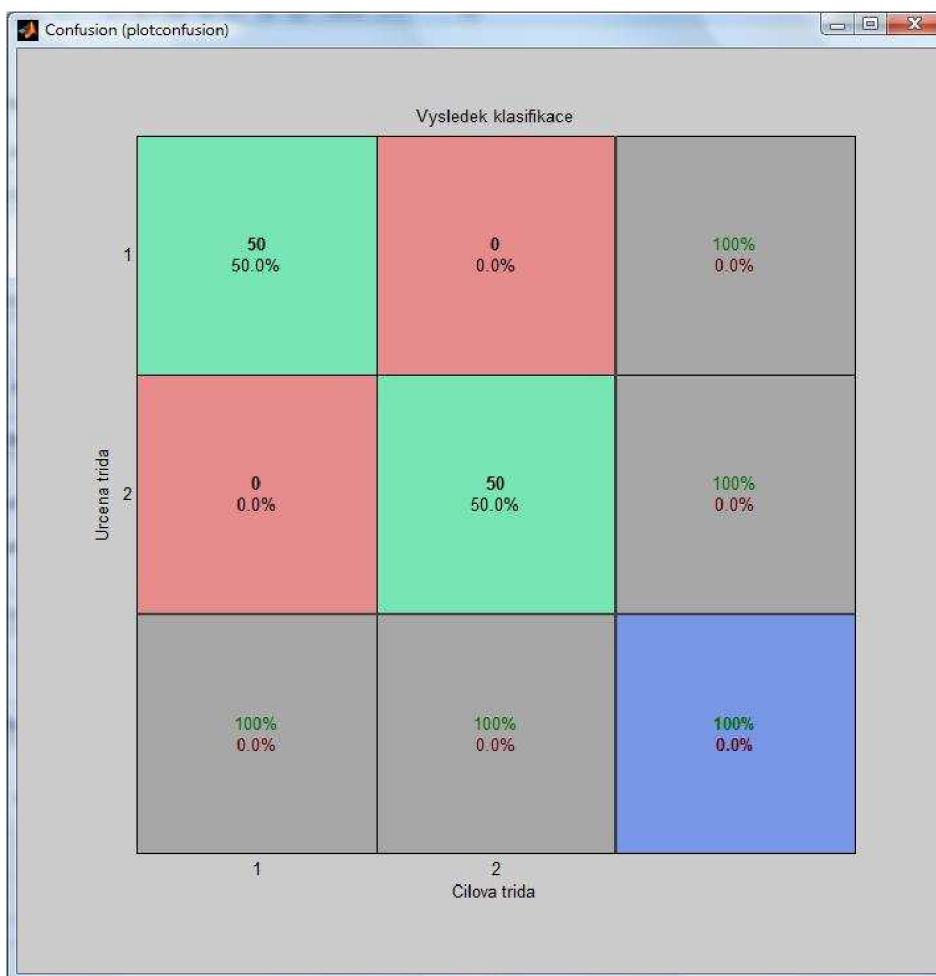
V této chvíli je nutné konkrétněji popsat navrženou neuronovou síť.

Pro dobrou funkci neuronové sítě je nutné zvolit vhodný počet skrytých vrstev v síti a optimální počet neuronů v těchto vrstvách. Proto byly provedeny experimenty, kdy se při stejných vstupních datech měnil tyto parametry a zkoumal se jejich vliv na výsledky klasifikace. Na základě výsledků těchto testů uvedených v kapitole 4.4 byla zvolena jedna vnitřní vrstva, počet neuronů nastaven na 10 a maximální počet epoch nastaven na 50. Dále byla nastavena přenosová funkce mezi vstupní a skrytou vrstvou, zvolena byla sigmoidní funkce, ta je v tomto případě nejpoužívanější. Také je nastavena přenosová funkce mezi skrytou vrstvou a výstupem, v tomto případě je zvolena jednoduchá lineární funkce. V případě použití dvou skrytých vrstev byla mezi nimi použita sigmoidní přenosová funkce. Dále byla zvolena metoda gradientního sestupu pro výpočet korekce vah a prahů. Následuje inicializace sítě a jak prahové hodnoty, taky i hodnoty vah jsou nastaveny na nuly. To proto, aby byly počáteční podmínky vždy stejné. Stiskem tlačítka *Klasifikovat data* bude spuštěno učení neuronové sítě pomocí trénovací matice (obr. 4.22).



Obr. 4.22: Učení neuronové sítě

Ukončení učení nastane ve třech případech, tj. buď bude dosažen maximální počet epoch – v tomto případě 50, nebo bude zpětně šířená chyba menší než nastavená hodnota 0,001, a nebo by hrozilo její „přeučení“, to znamená, že dalším učením by se zhoršila spolehlivost. Po naučení NS bude pomocí ní otestována testovací matice. Výsledek bude zobrazen pomocí funkce *plotconfusion*, která porovná výsledky testovací matice s přednastaveným vektorem a zjistí, kolik dat bylo zařazeno správně (obr. 4.23).



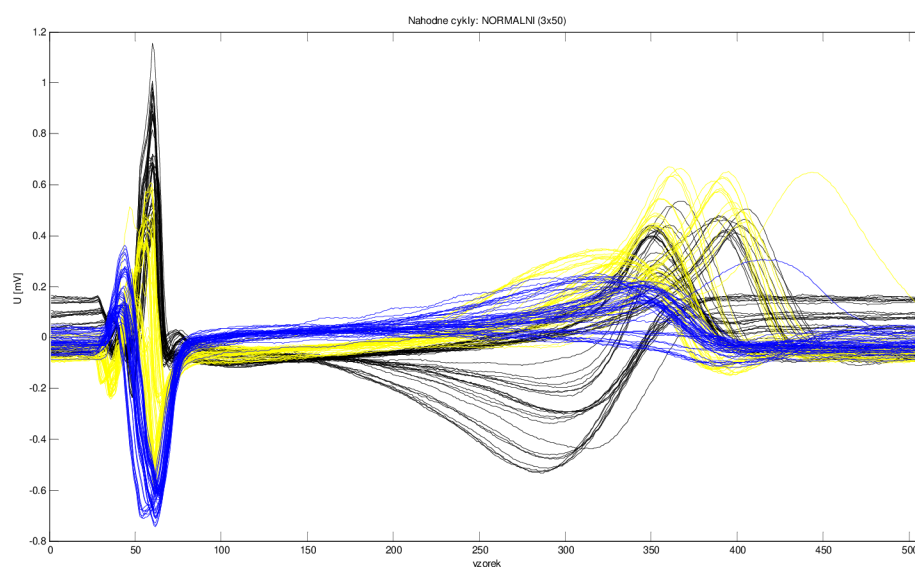
Obr. 4.23: Výsledek klasifikace

4.5.2 Aplikace s použitím PCA na všechny tři svody

Pomocí této aplikace jsou hlavní komponenty získávány vždy ze tří cyklů. Lépe řečeno ze tří průběhů, protože jeden srdeční cyklus byl monitorován pomocí tří svodů a byly tedy získány tři průběhy, které popisují jeden srdeční cyklus. Tyto tři průběhy jsou

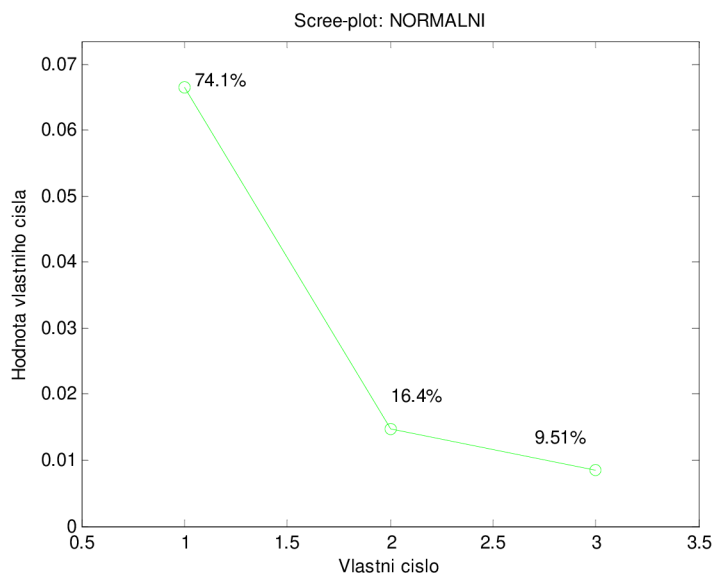
vstupem do PCA, výsledná hlavní komponenta tedy ponese informaci o jednom srdečním cyklu získanou ze tří signálů (svodů).

Spuštění aplikace se provádí otevřením souboru *triSVODY.m* a stiskem *run*. Podobně jako v předchozím případě je zde tlačítko *Nacist data*, ale už zde není panel pro výběr svodu, protože se pracuje se všemi třemi. Po stisknutí tlačítka *Nahodne vybrat cykly* bude náhodně vybráno sto padesát průběhů reprezentujících padesát srdečních cyklů pomocí tří svodů, tedy 3 x 50 křivek pro každou fázi. Do prvního grafu se zobrazí průběhy z normální fáze a do druhého grafu průběhy z ischemické fáze. Ukázka pro normální fázi je na obrázku 4.24.



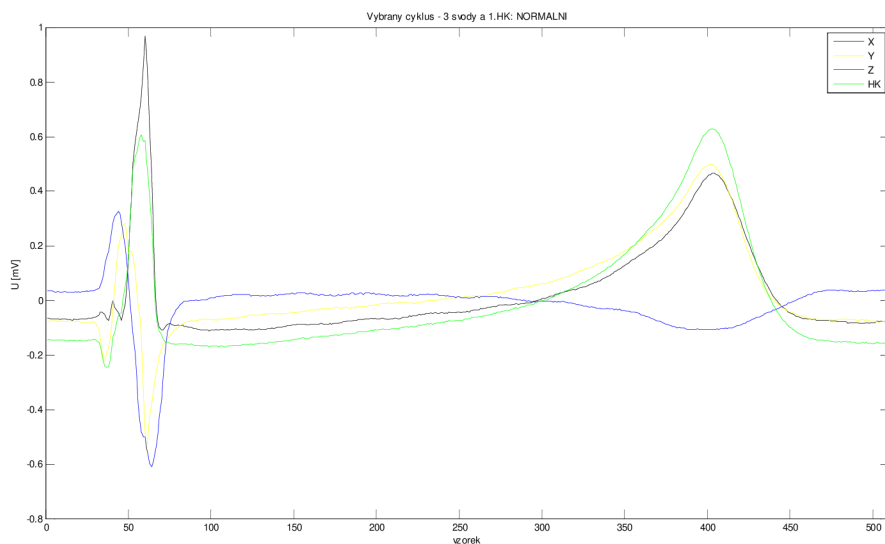
Obr. 4.24: Náhodně vybrané signály, svody X, Y a Z (3x50)

Jak už bylo řečeno, vstupem do jedné PCA budou tři průběhy a výstupem jeden průběh, který bude reprezentovat jednu, dvě nebo tři hlavní komponenty. Jelikož analýza proběhne celkem padesátkrát pro každou fázi, bude získáno padesát výstupních průběhů reprezentovaných hlavními komponentami pro každou fázi. Do klasifikace tedy bude vstupovat padesát průběhů z normální a padesát z ischemické fáze. Nyní je možné zobrazit scree ploty pro obě fáze pomocí příslušného tlačítka, na jejichž základě je možné zvolit vhodný počet komponent. Na obrázku 4.25 je scree plot pro normální fázi.



Obr. 4.25: Scree plot pro normální fázi

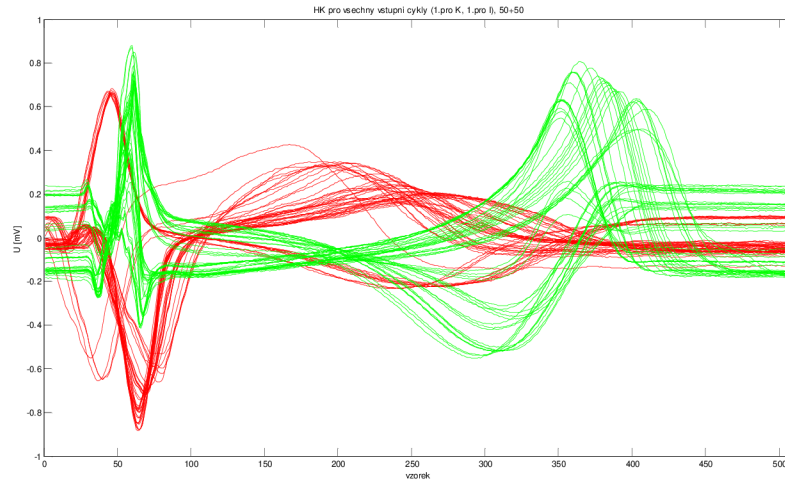
Po stisku dalšího tlačítka s názvem *Výsledky PCA* dojde k zobrazení ukázkových případů, do pátého grafu pro normální fázi a do šestého grafu pro ischemickou. Zobrazí se celkem čtyři průběhy, tři z nich jsou vstupy do PCA: svody X (černě), Y (žlutě) a Z (modře). Čtvrtý průběh je výsledek analýzy, tedy hlavní komponenta. Ukázkový případ pro normální fázi je na obrázku 4.26.



Obr. 4.26: Tři vstupní signály (svody X, Y, Z) a výsledná komponenta

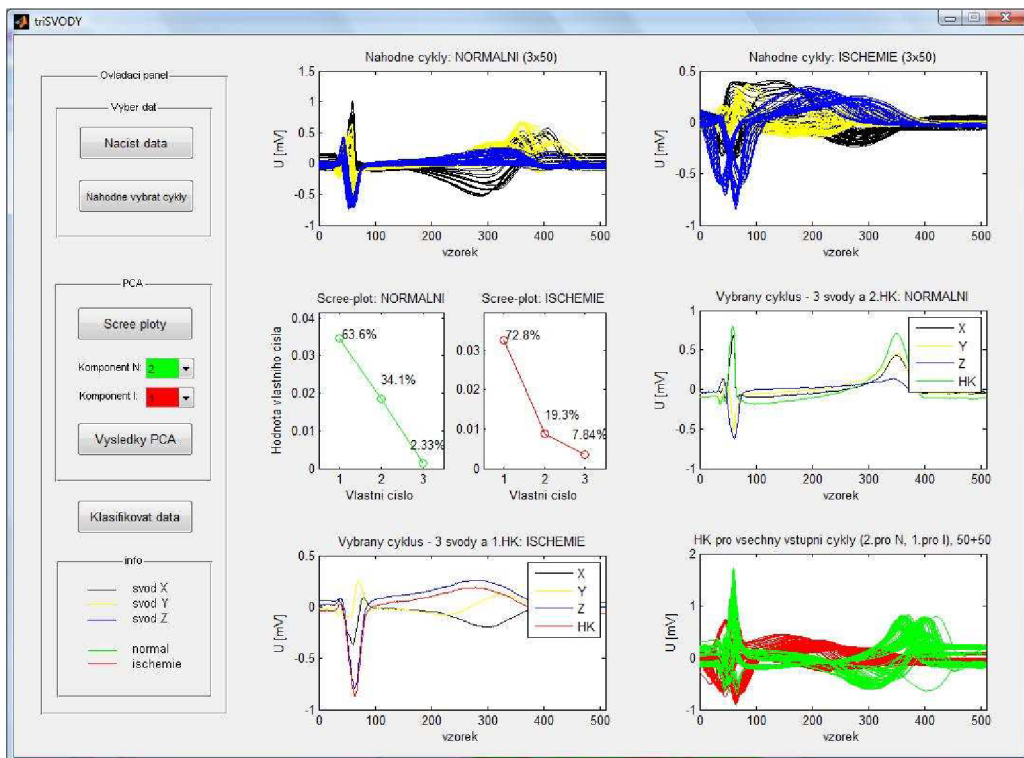
V posledním grafu budou zobrazeny hlavní komponenty ze všech provedených PCA,

celkem tedy padesát normálních a padesát ischemických (obr. 4.27).



Obr. 4.27: Výsledné hlavní komponenty ze tří svodů

Stejně jako v předchozím případě (kapitola 4.5.1) i nyní celý výše popsaný postup probíhal současně ještě jednou, aby byly získány dvě množiny dat. Opět jedna z nich bude sloužit k naučení neuronové sítě a druhá skupina dat bude na naučené neuronové síti testována. Rovněž byly v aplikaci zobrazeny pouze data jedné množiny, z důvodu lepší přehlednosti (obr. 4.28).



Obr. 4.28: Aplikace Tři svody

Znovu tedy budou vytvořeny dvě matice, jedna trénovací a druhá testovací. Nyní už zbývá jen stisknout tlačítko *Klasifikovat data*, kterým bude spuštěno učení neuronové sítě a následné testování dat. Podle výsledků testů uvedených v kapitole 4.4 byla nastavena jedna skrytá vrstva s devíti neurony a maximálním počtem 300 epoch při trénování sítě. Ostatní parametry sítě jsou stejné jako u aplikace popsané v kapitole 4.5.1.

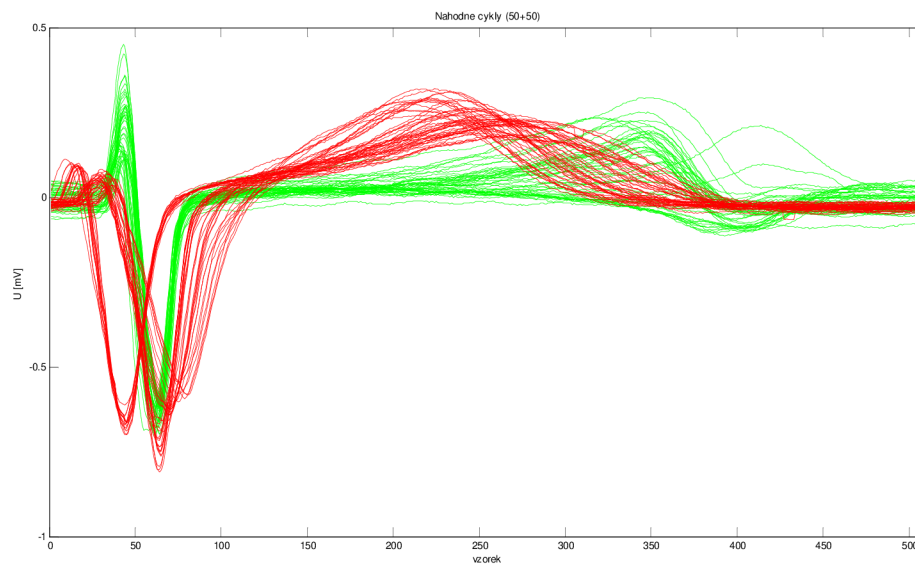
4.5.3 Aplikace s využitím redukce dat v rámci jednoho cyklu

V kapitole 4.3.3 již bylo popsáno, jakým způsobem budou komponenty získány a jak bude volen jejich počet. Níže je popsán konkrétní postup pro 50 náhodně vybraných vstupních cyklů a 20 zvolených komponent.

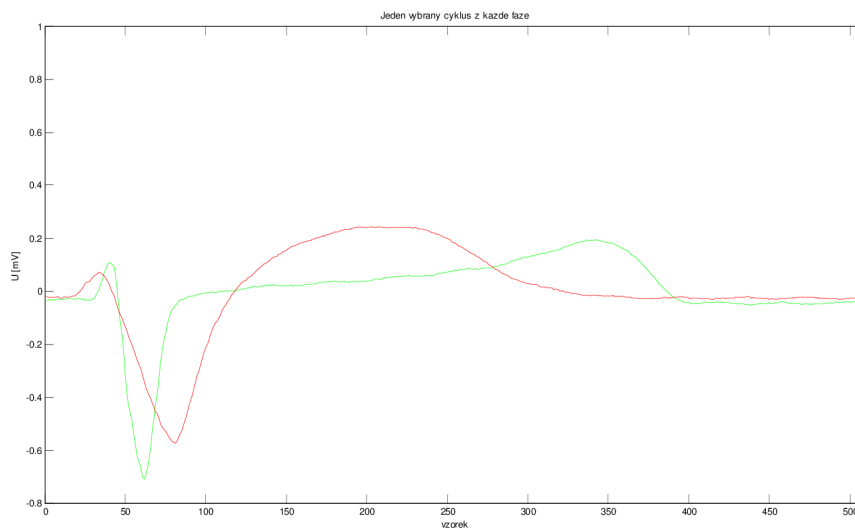
Bude náhodně vybráno 50 cyklů (pro každou fázi), vstupní matice má tedy 510 řádků a 50 sloupců. Matice ale musí být transponována kvůli tomu, aby bylo později možné provést násobení s jinou maticí. Tím se tedy rozměry upraví na 50 x 510. Z těchto hodnot budou vypočteny vlastní čísla a matice koeficientů (vlastních vektorů). Matice koeficientů je čtvercová, má velikost 510 x 510, přičemž sloupce odpovídají vlastním vektorům. Podle korelačního koeficientu mezi původním a rekonstruovaným cyklem bude zvolen požadovaný počet komponent, v tomto případě 20. Nyní bude vybráno z matice koeficientů tolik sloupců, kolik bylo zvoleno komponent. Tím dojde ke zredukování matice koeficientů, bude mít rozměry 510 x 20.

Nyní jsou tedy k dispozici dvě matice, vstupní o rozměrech 50 x 510 a matice vlastních vektorů s rozměry 510 x 20. V tomto okamžiku je třeba vynásobit vstupní cykly s maticí koeficientů a právě kvůli tomu byla vstupní matice transponována. Vstupní cykly však budou násobeny po jednom, takže v jednom kroku se bude jeden cyklus, což je vektor o rozměrech 1 x 510 násobit maticí 510 x 20. Výsledkem bude vektor 1 x 20, tedy dvacet hodnot, což bude prvních dvacet komponent (jeden cyklus bude reprezentován dvaceti hodnotami). Tímto způsobem budou vynásobeny všechny vstupní cykly, celkem jich tedy bude 50 pro každou fázi. Zpětně rekonstruovaný signál bude získán vynásobením získaného vektoru 1 x 20 se zredukovanou maticí vlastních vektorů, kterou je nutné transponovat – ze stejného důvodu, jako v prvním případě. Bude mít rozměry 20 x 510 a po vynásobení bude získán rekonstruovaný signál – vektor 1 x 510.

Po otevření souboru *redukce.m* a jeho spuštění dojde k otevření aplikace. Nejprve je opět nutné načíst data pomocí příslušného tlačítka. Následuje výběr svodu a náhodný výběr cyklů. Do prvního grafu budou zobrazeny náhodně vybrané cykly (50 z každé fáze, obr. 4.29) a do druhého grafu se vykreslí jeden průběh z každé fáze jako demonstrativní příklad (obr. 4.30).

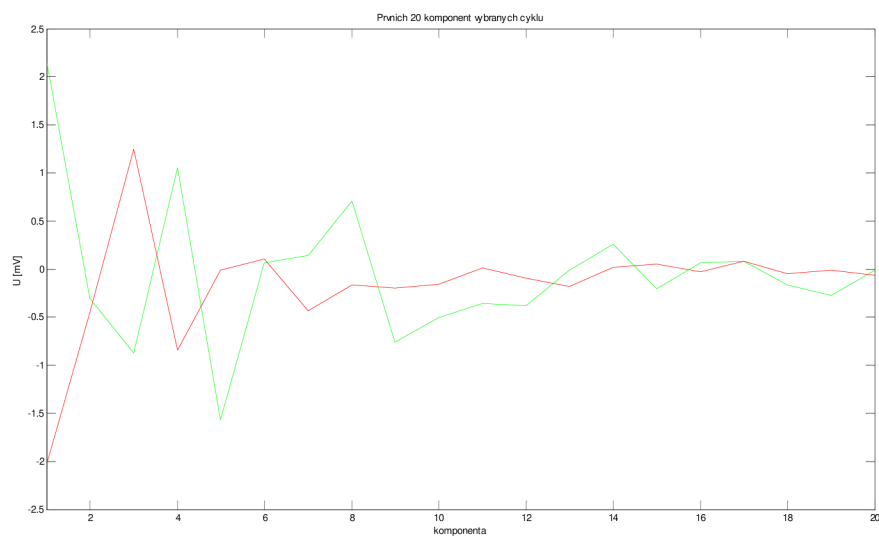


Obr. 4.29: Náhodně vybrané cykly z normální a ischemické fáze (50 + 50), svod Z

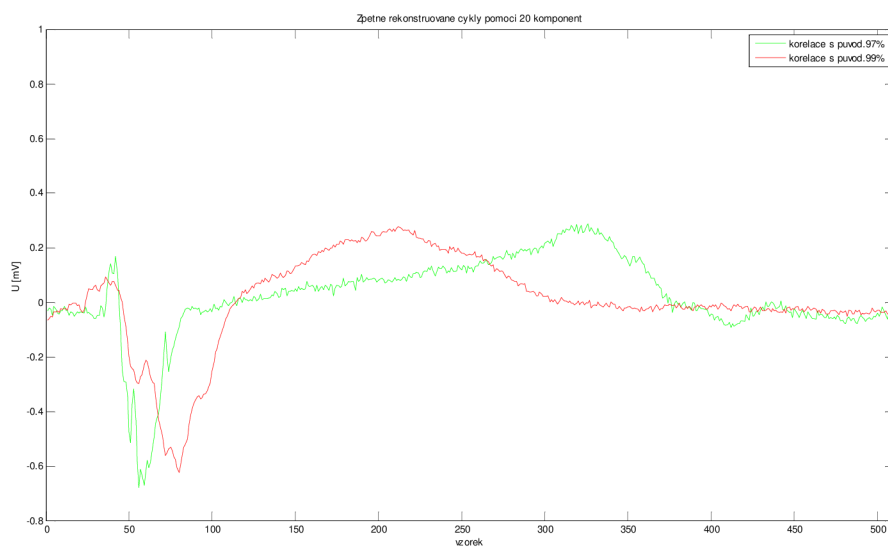


Obr. 4.30: Jeden vybraný cyklus z každé fáze, svod Z

Nyní je možné zvolit počet komponent, je zde přednastavená hodnota 20. V tomto okamžiku je vhodné stisknout tlačítko *Provest PCA* a zjistit, jak korelují vybrané cykly (obr. 4.30) s rekonstruovanými (obr. 4.32). Počet komponent je tedy možné upravit podle objektivního (korelace) i subjektivního (optické porovnání původního a zpětně rekonstruovaného signálu) kritéria. Ovšem korelace mezi původním a rekonstruovaným signálem by měla být přibližně 99%, kvůli následné spolehlivější klasifikaci. Do třetího grafu bude zobrazeno prvních 20 komponent (obr. 4.31) a do čtvrtého grafu rekonstruované cykly z těchto komponent (obr. 4.32).

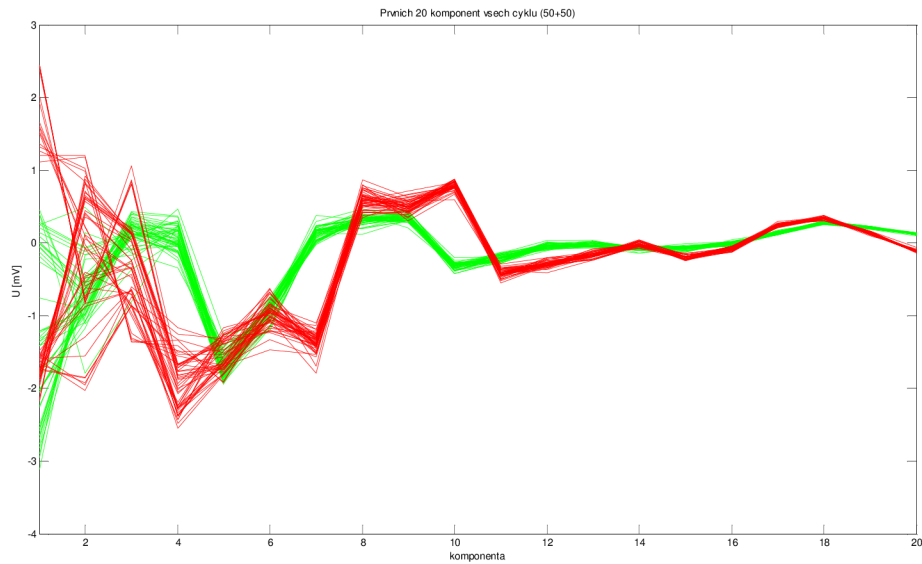


Obr. 4.31: Prvních 20 komponent vybraných cyklů

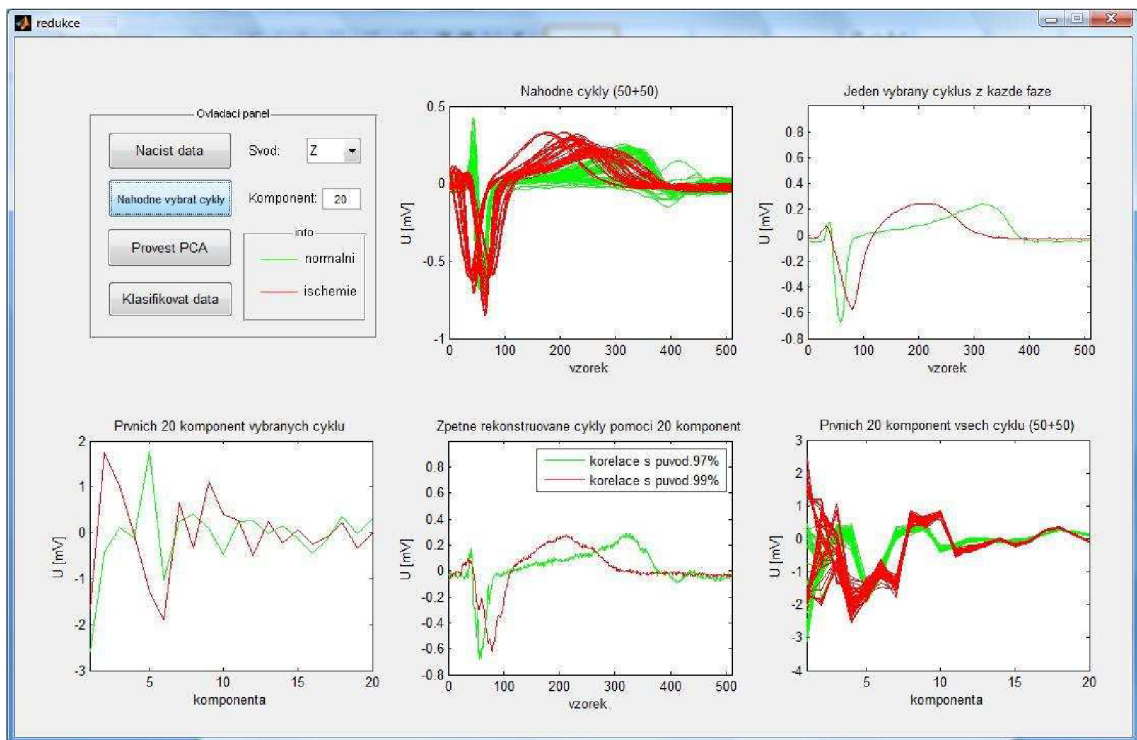


Obr. 4.32: Rekonstruované cykly pomocí 20-ti komponent

V posledním grafu bude možné vidět prvních 20 komponent pro všechny cykly (50 + 50), obrázek 4.33. Celá aplikace je pak na obrázku 4.34.



Obr. 4.33: Prvních 20 komponent všech cyklů



Obr. 4.34: Aplikace Redukce

Teď už zbývá data jen klasifikovat. Stejně jako v předchozích případech i teď celý proces probíhal paralelně ještě jednou a byly tak získány dvě množiny dat. Opět jedna bude sloužit pro naučení neuronové sítě a druhá pro testování.

Opět byly nastaveny nejhodnější parametry na základě výsledků testů popsaných v kapitole 4.4. Zvolena byla jedna skrytá vrstva s deseti neurony a maximálním počtem 300 epoch pro učení sítě.

4.6 Výsledky klasifikace QRST segmentů

V této části jsou uvedeny výsledky klasifikace. Jelikož vždy byly cykly vstupující do analýzy hlavních komponent vybírány náhodně, byla klasifikace provedena několikrát. Právě kvůli náhodnému výběru nebude vždy výsledek klasifikace stejný, ale bude se pohybovat v nějakém rozmezí. Jsou zkoumány výsledky pro různé svody i pro různou vyčerpanou variabilitu dat zvolením různého počtu komponent. V políčku vyčerpaná variabilita je uvedena přibližná variabilita dat z obou fází. Např. údaj 70% udává, že v normální fázi bylo vyčerpano 70% variability ($\pm 5\%$), v ischemické fázi rovněž 70% variability ($\pm 5\%$). Ve všech případech byly testovány dvě různé variability, menší (60 – 75%) a větší (80 – 95%). Menší variabilita obvykle odpovídala jedné komponentě, větší dvěma nebo třem komponentám. Výsledky jsou okomentovány na konci této kapitoly.

Jeden reprezentant ze sta QRST segmentů

Tab. 4.10: Úspěšnost klasifikace pro 1 ze 100, svod X

Svod X	Číslo pokusu	1	2	3	4	5	KLASIFIKACE		
	Vyčerpaná variabilita[%]	60	65	68	69	63	min	max	průměr
	Výsledek klasifikace[%]	100	100	100	100	100	100	100	100%
Svod X	Číslo pokusu	1	2	3	4	5	KLASIFIKACE		
	Vyčerpaná variabilita[%]	80	81	83	89	91	min	max	průměr
	Výsledek klasifikace[%]	100	100	100	100	100	100	100	100%

Tab. 4.11: Úspěšnost klasifikace pro 1 ze 100, svod Y

Svod Y	Číslo pokusu	1	2	3	4	5	KLASIFIKACE		
	Vyčerpaná variabilita[%]	71	62	72	58	65	min	max	průměr
	Výsledek klasifikace[%]	100	100	100	100	100	100	100	100%
Svod Y	Číslo pokusu	1	2	3	4	5	KLASIFIKACE		
	Vyčerpaná variabilita[%]	82	79	82	83	80	min	max	průměr
	Výsledek klasifikace[%]	100	100	100	100	100	100	100	100%

Tab. 4.12: Úspěšnost klasifikace pro 1 ze 100, svod Z

Svod Z	Číslo pokusu	1	2	3	4	5	KLASIFIKACE		
	Vyčerpaná variabilita[%]	81	79	79	78	81	min	max	průměr
	Výsledek klasifikace[%]	100	100	100	100	100	100	100	100%
Svod Z	Číslo pokusu	1	2	3	4	5	KLASIFIKACE		
	Vyčerpaná variabilita[%]	95	92	93	95	93	min	max	průměr
	Výsledek klasifikace[%]	100	100	100	100	100	100	100	100%

Jeden reprezentant ze tří svodů

Tab. 4.13: Úspěšnost klasifikace, tři svody

Svody X, Y, Z	Číslo pokusu	1	2	3	4	5	KLASIFIKACE		
	Vyčerpaná variabilita[%]	65	65	67	68	72	min	max	průměr
	Výsledek klasifikace[%]	94	99	95	99	97	94	99	96,8%
Svody X, Y, Z	Číslo pokusu	1	2	3	4	5	KLASIFIKACE		
	Vyčerpaná variabilita[%]	96	97	91	94	93	min	max	průměr
	Výsledek klasifikace[%]	96	96	92	91	97	92	97	94,6%

Redukce v rámci jednoho cyklu

V tomto případě byla kolonka *Vyčerpaná variabilita* nahrazena korelací mezi původním a rekonstruovaným cyklem.

Tab. 4.14: Úspěšnost klasifikace, svod X

Svod X	Číslo pokusu	1	2	3	4	5	KLASIFIKACE		
	Korelace[%]	88	87	80	90	91	min	max	průměr
	Výsledek klasifikace[%]	99	50	67	62	46	50	99	64,8%
Svod X	Číslo pokusu	1	2	3	4	5	KLASIFIKACE		
	Korelace[%]	99	99	99	99	99	min	max	průměr
	Výsledek klasifikace[%]	63	53	99	52	88	52	99	71%

Tab. 4.15: Úspěšnost klasifikace, svod Y

Svod Y	Číslo pokusu	1	2	3	4	5	KLASIFIKACE		
	Korelace[%]	89	91	91	88	87	min	max	průměr
	Výsledek klasifikace[%]	50	71	72	91	41	41	91	65%
Svod Y	Číslo pokusu	1	2	3	4	5	KLASIFIKACE		
	Korelace[%]	99	99	99	99	99	min	max	průměr
	Výsledek klasifikace[%]	51	62	73	57	100	51	100	68,6%

Tab. 4.16: Úspěšnost klasifikace, svod Z

Svod Z	Číslo pokusu	1	2	3	4	5	KLASIFIKACE		
	Korelace[%]	80	75	74	81	78	min	max	průměr
	Výsledek klasifikace[%]	42	66	99	52	68	42	99	65,4%
Svod Z	Číslo pokusu	1	2	3	4	5	KLASIFIKACE		
	Korelace[%]	99	99	99	99	99	min	max	průměr
	Výsledek klasifikace[%]	87	100	52	58	56	51	100	70,6%

V prvním případě, kdy byl jeden výsledný průběh (reprezentovaný komponentami)

získán ze sta vstupních cyklů, byla klasifikace vždy stoprocentní. Nezáleželo na tom, o který svod se jednalo, ani na vyčerpané variabilitě, vždy bylo dosaženo 100%. Lze tedy říct, že v tomto případě stačilo zvolit pouze první komponentu, která ponese přibližně 70% informace ze všech vstupních dat a bude to dostačující. Každopádně nebývá obvyklé, aby klasifikace byla úplně bezchybná. V tomto případě to může být dáno tím, že když mezi sto vstupními cykly byly některé výrazně odlišné od ostatních, tak byly pomocí metody hlavních komponent odstraněny. PCA je jednoduše nepovažovala za podstatné a zařadila je mezi nedůležitou informaci, která v tomto případě představuje 30%. Metodou hlavních komponent byly tedy získány průběhy, které si byly velmi podobné, a proto byla klasifikace stoprocentní.

Když byl klasifikován jeden srdeční cyklus popsán pomocí tří průběhů získaných ze tří svodů, výsledky klasifikace se pohybovaly lehce pod sto procenty. Konkrétně při zvolené jedné komponentě (což odpovídá variabilitě asi 68%) měla výsledná klasifikace hodnotu 96,8%. Hodnoty jednotlivých klasifikací se pohybují v rozmezí $\pm 5\%$. Pokud byla vyčerpana variabilita přibližně 95% (to odpovídá dvěma nebo třem komponentám) průměrná hodnota klasifikace byla rovna 94,6%. Výsledky jednotlivých klasifikací se pohybují rovněž v rozmezí $\pm 5\%$, ani zde tedy nejsou žádné velké výkyvy. Je však zajímavé, že při zvolení většího počtu komponent je klasifikace mírně horší. Může to být dáno tím, že větší vyčerpaná variabilita přináší více nepodstatné informace, tedy více šumu. Ovšem hodnoty klasifikace v obou případech jsou velmi dobré.

V posledním případě jsou výsledky klasifikace nejhorší. Pokud byla korelace mezi původním a rekonstruovaným signálem 99% (při volbě 20-ti komponent) výsledky měly lepší hodnoty, než když míra korelace odpovídala 90% (volba 5-ti komponent). Ukázalo se, že korelace má vliv na výslednou klasifikaci. Konkrétně při nižší korelaci byla klasifikace úspěšná průměrně pro 64,8% případů u svodu X, 65% u svodu Y, a 65,4% pro Z. U vyšší míry korelace to bylo 71% pro svod X, 68,6% pro Y a 70,6% pro svod Z. Z těchto výsledků je vidět, že nezáleží na volbě svodu, hodnoty pro všechny tři jsou velice podobné. Mezi výsledky jednotlivých klasifikací jsou velké rozdíly (až 50%), což může být dáno tím, že v tomto případě je klasifikovaný cyklus popsán pouze 20-ti hodnotami oproti předchozím dvěma případům, nebo také malým počtem trénovacích dat. To může mít za následek celkově horší klasifikaci. Lze tedy říci, že při vhodné volbě komponent (a odpovídající korelaci) jsou tato data klasifikována s úspěšností přibližně 70%.

5 ZÁVĚR

Dle zadání byl objasněn obecný princip metody hlavních komponent (PCA) a také její využití při analýze a následné klasifikaci srdečních cyklů. Ke zpracování dat byly zvoleny tři různé postupy zpracování QRST segmentů (reprezentujících srdeční cykly) využívající PCA. Cílem bylo zpracovat signály, což obnášelo vybrat podstatnou informaci (zvolit vhodný počet komponent) a snížit počet dat. Následně provést klasifikaci, která spočívala v rozřídění signálů na data z normální fáze experimentu a fáze předpokládané ischemie.

Ve všech třech případech byly využity signály z databáze dostupné na ÚBMI. V prvním případě bylo vstupem do analýzy sto QRST segmentů a výstupem jeden průběh reprezentovaný hlavními komponentami. Ve druhém případě vstupovaly do PCA tři průběhy (signál ze svodu X, svodu Y a svodu Z) reprezentující jeden QRST segment a výsledkem byl opět jeden průběh. V prvním i druhém případě tedy došlo ke snížení počtu dat: v prvním ze sta cyklů na jeden a ve druhém ze tří cyklů na jeden, přičemž všechny cykly měly stejný počet vzorků (510 vzorků). Třetí případ se liší od předchozích dvou. Zde se nesnižovala data tím, že by klesal počet cyklů, ale redukoval se počet vzorků popisující cyklus, konkrétně ze zmíněných 510 na 20 hodnot.

Pro každý ze tří zmíněných případů byla vytvořena speciální aplikace v programu MATLAB, díky které mohou být uživatelem snadno prováděny všechny úkony od výběru vstupních cyklů přes provedení PCA až po výslednou klasifikaci, která byla provedena pomocí vícevrstvé neuronové sítě.

Výsledky klasifikace u provedených případů jsou následující: Při stovce vstupních cyklů a jednom výstupním průběhu jsou všechna data klasifikována správně, klasifikace je tedy 100% pro všechny svody nezávisle na zvoleném počtu komponent. Tato bezchybná klasifikace může být dána tím, že pokud byly mezi vstupními cykly některé naprosto odlišné od ostatních, tak nebyly metodou PCA považovány za důležité, byly vyhodnoceny jako nepodstatná informace a tím byly odlišné cykly odstraněny. Výsledky druhého případu, kdy byl srdeční cyklus popsán signály ze třech svodů, už nejsou stoprocentní. Při takovém nastavení, kdy výstup PCA byl dán pouze jednou komponentou, spolehlivost klasifikace byla 96,8%. Při volbě dvou až tří komponent byla úspěšnost klasifikace 94,6%. Je zajímavé, že při volbě většího počtu komponent, tedy i při vyšší míře informace ze vstupních dat úspěšnost klasifikace mírně klesla. Je možné, že při vyšší míře informace ze vstupních dat stoupá i obsah šumu v těchto datech, a proto došlo k mírnému poklesu. To ale není příliš podstatné, výsledky této

části lze označit za více než uspokojivé. V posledním případě, kdy se snižoval počet vzorků jednotlivých cyklů, jsou výsledky klasifikace nejhorší. Při vhodném počtu komponent (tzn., že počet vzorků byl zredukován natolik, že mohl být zpětně obnoven s přesností na 99%) je úspěšnost klasifikace přibližně 70%. Toto číslo není příliš vysoké. Je možné, že nízkou míru klasifikace způsobuje malý počet vybraných komponent nebo malý počet trénovacích dat.

V porovnání výsledků s jinými autory jsou získané výsledky podobné. Je důležité zmínit, že u většiny analýz popsané jinými autory jsou vstupem do PCA nějaké vybrané parametry ze srdečního cyklu, neboli vektory příznaků. V této práci jsou na vstup PCA vždy dávány celé naměřené (předzpracované) signály. Ovšem výsledky jsou podobné jako u jiných autorů, u kterých je úspěšnost klasifikace v rozmezí 87 – 99%. To víceméně odpovídá dosaženým výsledkům v této práci, až na poslední případ, kdy byla klasifikace sedmdesátiprocentní.

LITERATURA

- [1] HAMPTON, R. John: *EKG stručně, jasně, přehledně*. Grada Publishing, Praha, 2005. ISBN 80-247-0960-0
- [2] HAMAN, P.: *Základy EKG*. Dostupné z: <<http://www.ekg.kvalitne.cz>>. [6.11.2012]
- [3] Smartdraw, dostupné z: <<http://www.smartdraw.com>>. [24.11.2012]
- [4] Sakshat Virtual Labs, dostupné z: <<http://iitr.vlab.co.in>>. [26.11.2012]
- [5] KOLÁTEK, J.: *Generátor patologických EKG křivek pro potřeby simulačních modelů*. Praha: ČVUT. Diplomová práce, 2010.
- [6] CASTELLS, F.; LAGUNA, P.; SÖRNMO, L.; BOLLMANN, A.; ROIG, J.: *Principal Component Analysis in ECG Signal Processing*. Vědecký článek, 2006.
- [7] CEYLAN, R.; ÖZBAY, Y.: *Comparasion of FCM, PCA and WT techniques for classification ECG arrhythmias using artificial neural network*. Konya, Selcuk University. Vědecký článek, 2007.
- [8] KOZUMPLÍK, J.; SMITAL, L.; VÍTEK, M.: *ELIMINATOR; Software pro korekci driftu nulové izolinie signálu EKG*. Ústav biomedicínského inženýrství, VUT v Brně, 2011. Dostupné z: <<http://www.ubmi.feec.vutbr.cz/vyzkum-a-vyvoj/produkty>>. [26.11.2012]
- [9] VÍTEK, M.; KOZUMPLÍK, J.: *QRS SEEKER; Software pro detekci komplexů QRS*. Ústav biomedicínského inženýrství, Vysoké učení technické v Brně, 2011.
- [10] JARKOVSKÝ, J.; LITTNEROVÁ, S.: *Pokročilé statistické metody*. Masarykova Univerzita Brno. Učební text, 2011.
- [11] JANKOWSKI, S.; DUSZA, J. J.; WIERZBOWSKI, M.; OREZIAK, A.: *PCA representation of ECG signal as a useful tool for detection of premature ventricular beats in 3-channel holter recording by neural network and support vector machine classifier*. Medical university of Warsaw, Poland. Vědecký článek, 2004.
- [12] IMAH, E. M.; AFIF, F. A.; FANANY M. I.; JATMIKO, W.; BASARUDDIN, T.: *A comparative study on daubechies zavelet transformation, Kernel PCA and PCA as feature extractors for arrhythmia detection using SVM*. Universitas Negeri, Surabaya. Vědecký článek, 2011.
- [13] HAMOUD, M. N.; MOUSSAOUI, A.: *Two novel methods for multiclass ECG arrhythmias classification based on PCA, fuzzy support vector machine and unbalanced clustering*. University of Cheick Larbi Tebessi, Algeria. Vědecký článek, 2010.
- [14] MURUGAN, S.; RADHAKRISHNAN, S.: *Improving ischemic beat classification using fuzzy genetic based PCA and ICA*. Department of ICE, Arulmigu Kalasalingam, India. Vědecký článek, 2010.
- [15] PALREDDY, S.; HU, Y. H.; MANI, V.; TOMPKINS, W. J.: *A multiple-clasiffier architecture for ecg beat classification*. IEEE Workshop on Neural Network for Signal Processing. Vědecký článek, 1997.
- [16] LINH, T. H.; OSOWKI, S.; STODOLSKI, M.: *Online heart beat recognition using hermite polynomials and neuro-fuzzy network*. IEEE Trans. Instrum. Meas. Vědecký článek, 2003.
- [17] PRASAD, G. K.; SAHAMBI, J. S.: *Classification of ecg arrhythmias using multi-resolution analysis and neural networks*. IEEE TECON. Vědecký článek, 2003.

- [18] CHAZAL, P.; O'DWYER M.; REILLY, R. B.: *Automatic classification of heartbeats using ecg morphology and heartbeat interval features*. IEEE Trans. Biomed. Eng. Vědecký článek, 2004.
- [19] RODRÍGUEZ, J.; GONI A.; ILLARRAMENDI A.: *Real time classification of ecgs on a pda*. IEEE Trans. Inform. Technol. Vědecký článek, 2005.
- [20] VONDRÁK, I.: *Neuronové sítě*. Fakulta elektrotechniky a informatiky VŠB. Učební text. Ostrava, 2009.
- [21] KANAAN, L.; MERHEB, D.; KALLAS, M.; FRANCIS, C.; AMOUD, H.; HONEINE, P.: *PCA and KPCA of ECG signals with binary SVM classification*. Université Saint Esprit de Kaslik, Jounieh, Lebanon. Vědecký článek, 2011.
- [22] KOLÁŘOVÁ, J.; FIALOVÁ, K.; JANOUŠEK, O.; NOVÁKOVÁ, M.; PROVAZNÍK, I. *Experimental methods for simultaneous measurement of action potentials and electrograms in isolated heart*. Physiological Research, 2010, ISSN: 0862- 8408.
- [23] PROCHÁZKA, T.: *Využití neuronových sítí pro klasifikaci alternací vlny T*. VUT v Brně. Diplomová práce, 2008.
- [24] ŠÍMA, J.; NERUDA, R.: *Teoretické otázky neuronových sítí*. Matfyzpress, Praha, 1996.
- [25] Posterus, dostupné z: <<http://www.posterus.sk>>. [15.4.2013]