

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

DIPLOMOVÁ PRÁCE

Brno, 2021

Bc. Klára Imramovská



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV BIOMEDICÍNSKÉHO INŽENÝRSTVÍ

DEPARTMENT OF BIOMEDICAL ENGINEERING

DETEKCE KOMOROVÝCH EXTRASYSTOL V EKG

PVC DETECTION IN ECG

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. Klára Imramovská

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Marina Ronzhina, Ph.D.

BRNO 2021

Diplomová práce

magisterský navazující studijní program **Biomedicínské inženýrství a bioinformatika**

Ústav biomedicínského inženýrství

Studentka: Bc. Klára Imramovská

ID: 195185

Ročník: 2

Akademický rok: 2020/21

NÁZEV TÉMATU:

Detekce komorových extrasystol v EKG

POKYNY PRO VYPRACOVÁNÍ:

1) Proveďte literární rešerši v oblasti projevů komorových extrasystol (KES) v EKG záznamech, kde popište i další patologie, které mohou mít podobné projevy v EKG (např. blokády Tawarových ramének aj.). Dále se seznamte s metodami pro automatickou detekci KES, využívajícími jednodušší postupy (template matching, rozhodovací pravidla s využitím EKG příznaků, „konvenční“ algoritmy strojového učení) i moderní techniky založené na hlubokém učení (CNN, autoencodery, aj.). Pro každou vybranou metodu uveďte v rešerši výhody/nevýhody. 2) Využijte veřejně dostupné databáze EKG s výskytem záznamů s KES a/nebo databázi patientských EKG záznamů dostupnou na ÚBMI. Seznamte se s databází a popište patologie vyskytující se v záznamech. Vytvořte dvě množiny EKG vhodné pro trénování a testování detekčních algoritmů. 3) Navrhněte metodu pro detekci KES a realizujte ji ve vhodném programovém prostředí. 4) Otestujte navrženou metodu na reálných záznamech. Vyhodnoťte úspěšnost použité metody. 5) Na základě předchozího bodu proveďte optimalizaci algoritmu pro dosažení co nejlepších výsledků. 6) Dosažené výsledky podrobně interpretujte. Například popište, ve kterých případech algoritmus chyboval a proč. Diskutujte výhody a limity navrženého detekčního postupu v kontextu jiných studií. Diskutujte možnosti vylepšení navrženého algoritmu.

DOPORUČENÁ LITERATURA:

[1] CAI, Z., et al. Rule-based rough-refined two-step-procedure for real-time premature beat detection in single-lead ECG. *Physiological Measurement* 2020, vol. 41, 054004, pp 1-21.

[2] GORDON, M. and WILLIAMS, C. PVC detection using a convolutional autoencoder and random forest classifier. *Pacific Symposium on Biocomputing* 2019, pp 42-53.

Termín zadání: 8.2.2021

Termín odevzdání: 21.5.2021

Vedoucí práce: Ing. Marina Ronzhina, Ph.D.

prof. Ing. Ivo Provazník, Ph.D.
předseda rady studijního programu

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

Práce se zabývá problematikou automatické detekce komorových extrasystol v EKG záznamech. V jazyce Python byla implementována metoda detekce využívající konvoluční neuronové sítě a LSTM jednotek. Pro detekci byly využity srdeční cykly extrahované z jednoho svodu EKG. Při klasifikaci do dvou tříd (KES a normální rytmus) dosáhlo F1 skóre na testovací množině 96,41 %, u klasifikace do tří tříd (KES, normální rytmus a ostatní arytmie) 81,76 %. V závěru práce je úspěšnost klasifikace zhodnocena a diskutována, dosažené výsledky pro klasifikaci do dvou tříd jsou srovnatelné s výsledky metod z jiných studií.

Klíčová slova

komorová extrasystola, elektrokardiogram, klasifikace, konvoluční neuronová síť, LSTM jednotka

Abstract

The thesis deals with problems of automatic detection of premature ventricular contractions in ECG records. One detection method which uses a convolutional neural network and LSTM units is implemented in the Python language. Cardiac cycles extracted from one-lead ECG were used for detection. F1 score for binary classification (PVC and normal beat) on the test dataset reached 96,41 % and 81,76 % for three-class classification (PVC, normal beat and other arrhythmias). Lastly, the accuracy of the classification is evaluated and discussed, the achieved results for binary classification are comparable to the results of methods described in different papers.

Keywords

premature ventricular contraction, electrocardiogram, classification, convolutional neural network, LSTM unit

Bibliografická citace

IMRAMOVSKÁ, Klára. *Detekce komorových extrasystol v EKG*. Brno, 2021. Dostupné také z: <https://www.vutbr.cz/studenti/zav-prace/detail/134417>. Diplomová práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav biomedicínského inženýrství. Vedoucí práce Ing. Marina Ronzhina, Ph.D.

Prohlášení autora o původnosti díla

Jméno a příjmení studenta:	<i>Klára Imramovská</i>
VUT ID studenta:	<i>195185</i>
Typ práce:	<i>Diplomová práce</i>
Akademický rok:	<i>2020/21</i>
Téma závěrečné práce:	<i>Detekce komorových extrasystol v EKG</i>

Prohlašuji, že svou závěrečnou práci jsem vypracovala samostatně pod vedením vedoucí závěrečné 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 autorka uvedené závěrečné práce dále prohlašuji, že v souvislosti s vytvořením této závěrečné práce jsem neporušila autorská práva třetích osob, zejména jsem nezasáhla nedovoleným způsobem do cizích autorských práv osobnostních a jsem si plně vědoma 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.

V Brně dne: 21. května 2021

podpis autorky

Poděkování

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

V Brně dne: 21. května 2021

podpis autorky

Obsah

SEZNAM OBRÁZKŮ	9
SEZNAM TABULEK	11
ÚVOD	12
1. ELEKTRICKÁ AKTIVITA SRDCE	13
1.1 PŘEVODNÍ SYSTÉM SRDEČNÍ	13
1.2 ELEKTROKARDIOGRAM	14
1.3 SRDEČNÍ EXTRASYSTOLY	16
1.3.1 Síňové extrasystoly	17
1.3.2 Komorové extrasystoly	17
1.4 RAMÉNKOVÉ BLOKÁDY	21
1.4.1 Blokáda pravého raménka Tawarova	22
1.4.2 Blokáda levého raménka Tawarova	23
2. METODY DETEKCE KOMOROVÝCH EXTRASYSTOL	24
2.1 METODY ZALOŽENÉ NA VÝBĚRU RELEVANTNÍCH PŘÍZNAKŮ	24
2.1.1 Metody založené na aplikaci rozhodovacích pravidel	24
2.1.2 Metody využívající časových, morfologických či statistických příznaků	25
2.1.3 Metody využívající analýzu hlavních komponent	26
2.1.4 Metody založené na vyhledávání vzorů	27
2.1.5 Metody využívající hluboké neuronové sítě	28
2.1.6 Metody založené na výpočtu korelace	29
2.2 METODY VYUŽÍVAJÍCÍ HLUBOKÉ UČENÍ	30
2.2.1 Metody využívající autoenkodéru	30
2.2.2 Metody využívající konvoluční neuronové sítě	32
3. HLUBOKÉ UČENÍ	33
3.1 KONVOLUČNÍ NEURONOVÉ SÍTĚ	33
3.1.1 Typy vrstev používaných v CNN	33
3.1.2 Aktivační funkce	36
3.1.3 Kriteriaální funkce	39
3.1.4 Optimalizační metody	40
3.2 LSTM	42
3.2.1 Zapomínací brána	43
3.2.2 Vstupní brána	43
3.2.3 Výstupní brána	44
4. REALIZACE AUTOMATICKÉ DETEKCE KOMOROVÝCH EXTRASYSTOL	45
4.1 POUŽITÁ TESTOVACÍ DATABÁZE	45
4.2 VYTVOŘENÍ TRÉNOVACÍ, VALIDAČNÍ A TESTOVACÍ MNOŽINY	46
4.3 REALIZACE KLASIFIKACE EKG Z ÁZNAMŮ DO DVOU TŘÍD	49
4.3.1 Architektura sítě	50
4.3.2 Výběr vhodných hyperparametrů	51
4.3.3 Trénování sítě	55
4.3.4 Ukázky výsledků klasifikace	58

4.4	KLASIFIKACE EKG ZÁZNAMŮ DO TŘÍ TŘÍD.....	61
5.	HODNOCENÍ ÚSPĚŠNOSTI KLASIFIKACE	70
5.1	POUŽITÉ PARAMETRY PRO HODNOCENÍ ÚSPĚŠNOSTI	70
5.2	HODNOCENÍ ÚSPĚŠNOSTI NAVRŽENÉHO KLASIFIKÁTORU	71
5.2.1	<i>Klasifikace do dvou tříd.....</i>	<i>71</i>
5.2.2	<i>Klasifikace do tří tříd</i>	<i>75</i>
5.3	POROVNÁNÍ DOSAŽENÝCH VÝSLEDKŮ S PUBLIKOVANÝMI METODAMI	78
5.4	DISKUSE DOSAŽENÝCH VÝSLEDKŮ	84
	ZÁVĚR.....	88
	LITERATURA.....	89
	SEZNAM SYMBOLŮ A ZKRATEK.....	96

SEZNAM OBRÁZKŮ

1.1	Převodní systém srdeční (převzato a upraveno z [7])	14
1.2	Fyziologická křivka EKG (převzato z [10])	15
1.3	Síňová extrasystola v druhém, čtvrtém, šestém a osmém srdečním úderu (převzato z [13]).....	17
1.4	Multifokální komorová extrasystola (převzato z [13])	18
1.5	Komorová bigeminie (převzato z [13]).....	18
1.6	Komorová trigeminie (převzato z [13])	19
1.7	Kuplet (převzato z [13])	19
1.8	Invertovaná vlna P viditelná za třetím QRS komplexem (převzato z [13]).....	19
1.9	Interpolovaná komorová extrasystola s prodlouženým následujícím PQ intervalem, kompenzační pauza chybí (převzato z [13])	20
1.10	Fenomén „R na T“ vyvolávající fibrilaci komor (převzato z [13])	20
1.11	End-diastolická komorová extrasystola (převzato z [13])	20
1.12	Blokáda pravého Tawarova raménka (převzato z [13]).....	22
1.13	Blokáda levého Tawarova raménka (převzato z [13])	23
2.1	QRS vzory (převzato a upraveno z [21])	25
2.2	Příznaky získané z EKG signálu (převzato a upraveno z [29])	29
3.1	Princip konvoluční vrstvy (převzato a upraveno z [39]).....	34
3.2	Příklad poolingů s filtrem velikosti 2x2 a délkou kroku 2 (převzato z [42]).....	35
3.3	Plně propojené vrstvy (převzato z [44]).....	36
3.4	Vlevo příklad skokové funkce (převzato z [45]), vpravo lineární funkce (převzato z [46]).....	37
3.5	Sigmoidální a tanh funkce (převzato z [46]).....	38
3.6	Vlevo ukázka ReLU, vpravo Leaky ReLU (převzato z [46]).....	39
3.7	Nahoře struktura LSTM, dole stav buňky (převzato z [57]).....	43
3.8	Zapomínací brána (převzato z [57]).....	43
3.9	Vstupní brána (převzato z [57]).....	44
3.10	Výpočet nového stavu buňky (převzato z [57]).....	44
3.11	Výstupní brána (převzato z [57]).....	44
4.1	Ukázka doplnění úseku normálního rytmu oběma způsoby: vlevo nahoře původní úsek, vpravo nahoře úsek po doplnění prvním způsobem, dole úsek po doplnění druhým způsobem (viz text)....	47
4.2	Ukázka doplnění úseku KES oběma způsoby: vlevo nahoře původní úsek, vpravo nahoře úsek po doplnění prvním způsobem, dole úsek po doplnění druhým způsobem (viz text)	48
4.3	Část kódu věnující se nastavení callback funkce.....	52
4.4	Část kódu věnující se nastavení trénování sítě	52
4.5	Ukázka kolísání chybové funkce a přeučení sítě, velikost dávky = 32.....	54
4.6	Ukázka přeučení sítě při vyšším počtu epoch.....	54
4.7	Ukázka trénování sítě při špatném nastavení hyperparametrů	55
4.8	Průběh chybové funkce při trénování sítě, data doplněná prvním způsobem	56
4.9	Průběh trénovací a validační přesnosti při trénování sítě, data doplněná prvním způsobem	56
4.10	Průběh chybové funkce při trénování sítě, data doplněná druhým způsobem.....	57
4.11	Průběh trénovací a validační přesnosti při trénování sítě, data doplněná druhým způsobem	57
4.12	Ukázka falešně pozitivní klasifikace segmentů (data doplněná prvním způsobem).....	59
4.13	Ukázka falešně negativní klasifikace segmentů (data doplněná prvním způsobem)	60
4.14	Ukázka falešně pozitivní klasifikace segmentů (data doplněná druhým způsobem).....	60
4.15	Ukázka falešně negativní klasifikace segmentů (data doplněná druhým způsobem)	61
4.16	Průběh chybové funkce při trénování sítě pro klasifikaci do tří tříd	64
4.17	Průběh trénovací a validační přesnosti při trénování sítě pro klasifikaci do tří tříd.....	64

4.18 Ukázka chybně klasifikovaných segmentů normálního rytmu.....	65
4.19 Ukázka chybně klasifikovaných segmentů KES.....	67
4.20 Ukázka chybně klasifikovaných segmentů ostatních arytmií: SVES	67
4.21 Ukázka chybně klasifikovaných segmentů ostatních arytmií: BLRT	68
4.22 Ukázka chybně klasifikovaných segmentů ostatních arytmií: BPRT	69

SEZNAM TABULEK

1.1	Lownova klasifikace komorových extrasystol	21
4.1	Rozdělení dat	48
4.2	Počty srdečních úderů v jednotlivých záznamech.....	49
4.3	Architektura sítě.....	50
4.4	Nastavení hyperparametrů sítě pro klasifikaci do dvou tříd	55
4.5	Rozdělení dat pro klasifikaci do tří kategorií.....	62
4.6	Počty srdečních úderů dalších arytmií v jednotlivých záznamech.....	62
4.7	Nastavení hyperparametrů sítě pro klasifikaci do tří tříd.....	63
5.1	Matice záměn pro trénovací data doplněná prvním způsobem	71
5.2	Matice záměn pro validační data doplněná prvním způsobem.....	71
5.3	Matice záměn pro testovací data doplněná prvním způsobem.....	72
5.4	Parametry hodnocení úspěšnosti klasifikace pro data doplněná prvním způsobem.....	72
5.5	Matice záměn pro trénovací data doplněná druhým způsobem.....	73
5.6	Matice záměn pro validační data doplněná druhým způsobem.....	73
5.7	Matice záměn pro testovací data doplněná druhým způsobem	74
5.8	Parametry hodnocení úspěšnosti klasifikace pro data doplněná druhým způsobem.....	74
5.9	Porovnání úspěšnosti klasifikace pro oba způsoby doplnění vzorků.....	75
5.10	Matice záměn pro trénovací data.....	75
5.11	Úspěšnost klasifikace trénovacích dat pro všechny tři třídy.....	76
5.12	Matice záměn pro validační data	76
5.13	Úspěšnost klasifikace validačních dat pro všechny tři třídy	77
5.14	Matice záměn pro testovací data	77
5.15	Úspěšnost klasifikace testovacích dat pro všechny tři třídy	78
5.16	Porovnání dosažených výsledků s publikovanými metodami	79

ÚVOD

Komorové extrasystoly neboli předčasné stahy srdečních komor představují jednu z nejčastějších komorových arytmií. V převážné většině případů jsou asymptomatické a nepředstavují žádné riziko, někdy ale mohou být projevem či dokonce původcem závažných srdečních onemocnění. Detekce srdečních arytmií představuje v současné době jeden z velmi důležitých úkolů, přesná a včasná detekce komorových extrasystol totiž může být vzhledem k riziku rozvinutí vážných srdečních onemocnění klíčová. Základní vyšetřovací metodou kardiovaskulárních nemocí je elektrokardiografie.

Správné vyhodnocení elektrokardiografického záznamu je časově náročné, závisí především na odborných znalostech a zkušenostech lékaře. Proto bylo navrženo velké množství metod automatické detekce komorových extrasystol, v poslední době především těch, které ke klasifikaci využívají hluboké učení. Cílem této práce je navržení metody automatické detekce komorových extrasystol využívající metod hlubokého učení.

Teoretická část práce je věnována elektrické aktivitě srdce, převodnímu systému srdečnímu a elektrokardiogramu. Dále se zabývá podstatou komorových extrasystol a raménkových blokád a jejich projevy v elektrokardiografickém záznamu. V kapitole 2 jsou popsány principy vybraných metod automatické detekce komorových extrasystol. V kapitole 3 jsou představeny dvě metody hlubokého učení využitě v praktické části, a to konkrétně konvoluční neuronové sítě a LSTM jednotky.

V praktické části práce je pomocí programového prostředí Matlab a programovacího jazyka Python s využitím programu Spyder a knihoven Keras a TensorFlow realizována metoda klasifikace komorových extrasystol využívající konvoluční neuronové sítě a LSTM jednotek, která je otestována na záznamech z MIT-BIH Arrhythmia Database ze serveru PhysioNet. Srdeční rytmus je nejprve klasifikován jako do dvou tříd jako normální rytmus nebo komorová extrasystola, poté je klasifikace rozšířena o ostatní arytmie, kam jsou zařazeny síňové extrasystoly a raménkové blokády. V kapitole 5 jsou výsledky automatické detekce zhodnoceny a porovnány s výsledky některých publikovaných metod. Na závěr jsou diskutovány dosažené výsledky, výhody, nevýhody a limitace navržené metody a její možná vylepšení či rozšíření.

1. ELEKTRICKÁ AKTIVITA SRDCE

V následující kapitole je popsán převodní srdeční systém spolu s fyziologickým elektrokardiogramem (EKG). Dále jsou představeny některé srdeční arytmie včetně komorových extrasystol (KES) a raménkových blokády.

1.1 Převodní systém srdeční

Buňky srdeční svaloviny se dají na základě jejich funkce rozdělit na převodní systém srdeční a pracovní myokard.

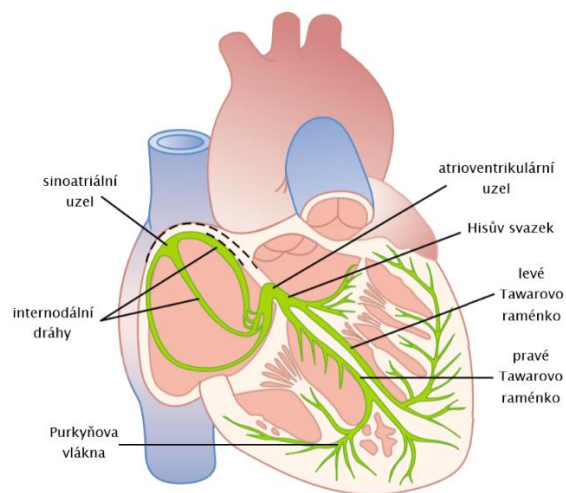
Převodní systém srdeční je specializovaný na elektrickou aktivitu. Má schopnost samovolně generovat elektrické impulsy a následně je rychle a koordinovaně rozvést po celém srdci. V buňkách převodního systému není konstantní klidový membránový potenciál, protože v nich probíhá spontánně pomalá depolarizace směrem ke spouštěcí úrovni akčního potenciálu. Po jeho skončení se hodnota membránového potenciálu nestabilizuje, ale opět se pomalu přibližuje spouštěcí úrovni dalšího akčního potenciálu. Jelikož se tento mechanismus neustále pravidelně opakuje, buňky převodního systému fungují jako zdroj vzruchů (pacemaker).

Buňky pracovního myokardu za normálních okolností nejsou schopny samovolné tvorby vzruchů (automacie), jejich hlavní funkce je reakce na elektrické impulsy kontrakcí srdečního svalu vedoucí k mechanické čerpací práci srdce. Akční potenciál těchto buněk je generován šířením vzruchu, k jeho vzniku stačí depolarizace o 20 až 25 mV [1],[2].

Buňky srdečního svalu jsou propojeny pomocí interkalárních disků. Jejich součástí jsou tzv. gap junctions složené z membránových kanálků, které umožňují snadný průchod iontů mezi buňkami. Buňky myokardu jsou tak spojeny do funkčního syncytia a pokud je excitována jedna buňka, šíří se akční potenciál na všechny ostatní. Srdce se prakticky skládá ze síňového a komorového syncytia [3],[4]. Převodní systém srdeční (viz obrázek 1.1) je tvořen sinoatriálním (SA) uzlem, internodálními síňovými spoji, atrioventrikulárním (AV) uzlem, Hisovým svazkem, Tawarovými raménky s větvením a Purkyňovými vlákny [1],[2].

V SA uzlu, který je umístěn pod epikardem ve stěně pravé síně blízko ústí horní duté žíly, je samovolná depolarizace nejrychlejší a nejdříve dosáhne spouštěcí úrovně. SA uzel proto funguje jako přirozený (primární) pacemaker, a frekvence jeho vzruchů tak určuje frekvenci srdečních stahů (přibližně 60–100/min) [5]. Rytmus vycházející z SA uzlu se označuje jako sinusový. Vzruch se odtud šíří po celém pracovním myokardu předsíní, přednostně ale využívá internodální dráhy, které vedou vzruch rychleji než pracovní myokard [6]. Vzruch postupně přechází k AV uzlu, který se nachází pod endokardem na spodině pravé předsíně. Tento uzel tvoří spolu s horní částí Hisova svazku tzv. AV junkci, která představuje jediné vodivé spojení mezi předsíněmi a komorami, spojuje tak obě

syncytia [3]. Síňokomorové rozhraní je totiž kromě ní tvořeno nevodivou vazivovou tkání srdečního skeletu. Velmi důležitou vlastností AV uzlu je pomalé vedení akčního potenciálu, jehož následkem je zdržení přenosu vzruchu z předsíní na komory dlouhé přibližně 100 ms. Kontrakce (systola) síní je tak oddělena od kontrakce komor, díky tomu srdce pumpuje krev s vysokou efektivitou [5]. Přes Hisův svazek a pravé Tawarovo raménko se vzruch dále dostává k myokardu pravé komory, levé raménko vede skrz mezikomorové septum k levé komoře. Vzruch je dalším větvením obou ramének doveden až do sítě Purkyňových vláken, která podráždění rozvádí na pracovní myokard komor zevnitř k vnějším vrstvám od hrotu směrem k bázi srdce [1],[2].



Obrázek 1.1 Převodní systém srdeční (převzato a upraveno z [7])

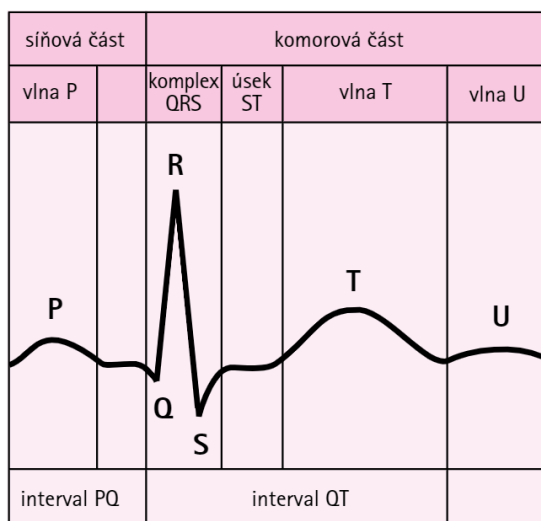
Rychlost šíření depolarizační vlny po pracovním myokardu se pohybuje od 0,3 do 1 m/s. V buňkách převodního systému srdečního v předsíních a komorách vzruch vede rychleji, šíří se rychlostí 1 až 4 m/s, nejrychlejší je v Purkyňových vlákních. Výjimkou jsou buňky SA a AV uzlu, ve kterých je rychlost vedení vzruchu pouze 0,02–0,1 m/s [1],[2].

Spontánní depolarizace je v dalších částech převodního systému pomalejší než v SA uzlu, a tak k nim dorazí vzruch z SA uzlu dříve, než jejich membránový potenciál dosáhne spouštěcí úrovně. Za určitých podmínek se však zdrojem depolarizace mohou stát i buňky mimo fyziologický primární pacemaker. Frekvence tvorby vzruchů je ale pomalejší [1],[2]. V případě přerušení převodu sinusových podnětů přebírá roli pacemakeru AV uzel (sekundární pacemaker) s frekvencí impulsů přibližně 40–55 za minutu. Takto generovanému srdečnímu rytmu se říká nodální (junkční) [5],[6]. Za některých okolností mohou být Purkyňovy vlákna tzv. terciárním pacemakerem, generují impulsy s frekvencí 25–40/min. Tento rytmus je označován jako idioventrikulární [3],[4].

1.2 Elektrokardiogram

Elektrokardiografie je jednou ze základních neinvazivních vyšetřovacích metod srdce, je

důležitým ukazatelem při diagnostice řady srdečních onemocnění [8]. V jednotlivých myokardiálních buňkách vzniká v průběhu akčního potenciálu, kdy je depolarizována pouze část buňky, elementární elektrické pole. Sumací elektrických projevů všech srdečních buněk vznikají změny elektrického napětí, které mohou být zaznamenány i na povrchu těla. Výsledná suma v daném okamžiku může být charakterizována okamžitým vektorem elektrického pole srdečního. Elektrokardiogram je grafický záznam elektrické aktivity srdce získaný pomocí elektrod přiložených na povrch těla [9]. Běžně se EKG záznam skládá ze 12 svodů rozdělených do tří skupin, podává informaci o srdečním rytmu, tepové frekvenci, původu a šíření vzruchů či o poloze srdce [1],[3],[4].



Obrázek 1.2 Fyziologická křivka EKG (převzato z [10])

EKG křivku tvoří časový průběh okamžitých sumačních vektorů elektrického pole, které jsou charakterizovány velikostí a směrem. Podle zavedené konvence je výchylka kladná, pokud směřuje vlna depolarizace směrem k elektrodě, naopak v případě depolarizační vlny směřující od elektrody je výchylka záporná. Pro repolarizaci platí toto pravidlo obráceně. V případě, kdy sumační vektor směřuje kolmo ke svodu, je výchylka nulová [4],[11]. Na EKG křivce, zobrazené na obrázku 1.2, můžeme rozeznat vlny (zaoblené výchylky s nižší amplitudou), kmity (úzké a hrotnaté výchylky), dále úseky neboli segmenty (vzdálenosti mezi vlnami a kmity) a intervaly (nejméně jedna vlna či kmit a úsek dohromady) [1],[12].

Vlna P zachycuje elektrickou aktivitu síní. Vzruch vychází z SA uzlu umístěného v pravé síni a depolarizační vlna se šíří skrz myokard síní. Pravá síň se začne depolarizovat dříve než levá, proto odpovídá první část P vlny převážně depolarizaci pravé síně a druhá část depolarizaci levé síně. P vlna trvá přibližně 80–120 ms a její amplituda je poměrně malá, protože stěna předsíní obsahuje oproti komorám méně svalové hmoty. Repolarizace síní není na EKG záznamu viditelná, jelikož je za normálních okolností překryta prominentnějšími kmity, které odpovídají depolarizaci komor [1],[6],[8].

Izoelektrický úsek PQ odpovídá zpomalení šíření vzruchu při převodu AV uzlem z předsíní na komory, kdy jsou síně kompletně depolarizovány. Normální PQ interval by měl trvat 120–200 ms [11],[12].

QRS komplex představuje depolarizaci komor. Elektrický impuls projde postupně Hisovým svazkem a oběma Tawarovými raménky skrz myokard mezikomorového septa až do Purkyňových vláken. Nejprve je depolarizováno mezikomorové septum, podráždění postupuje ve směru od levé komory k pravé. Vzruch se šíří od srdečního hrotu po pracovní svalovinu obou komor. Vzhled QRS komplexu může být variabilní, neboť odpovídá postupu šíření aktivace přes obě komory. První negativní výchylka je označena jako Q vlna, první pozitivní výchylka jako vlna R. S vlna je první negativní výchylka po vlně R. Případná další pozitivní výchylka je označena apostrofem jako R' [8]. Výchylky s amplitudou vyšší než 0,5 mV jsou označeny velkým písmenem, menší výchylky se značí malým písmenem. Amplituda QRS komplexu je vyšší než u vlny P, svalovina komor je totiž mnohem mohutnější než svalovina předsíní. Většina QRS komplexu představuje aktivaci levé komory, protože je svalová hmota levé komory přibližně třikrát silnější než svalovina komory pravé. Délka trvání fyziologického komplexu QRS je 60 až 100 ms [1],[4],[9].

Izoelektrická linie úseku ST mezi koncem QRS komplexu a začátkem vlny T reprezentuje fázi plató, kdy je elektrická aktivita srdce nulová, srdeční buňky mají stejný elektrický náboj a depolarizace je kompletně rozšířená po obou komorách. ST úsek odpovídá kontrakci komor, představuje čas mezi depolarizací a repolarizací komor.

Vlna T odráží repolarizaci komorového myokardu. Je širší než QRS komplex, trvá zhruba 0,2 s, jelikož je repolarizace komor pomalejší než jejich depolarizace.

Interval QT, který zaznamenává celkovou délku trvání depolarizace a repolarizace komor, závisí na věku, pohlaví, a hlavně srdeční frekvenci. S rostoucí srdeční frekvencí se zkracuje, při tepové frekvenci 75 úderů za minutu je dlouhý 0,35–0,40 s [11],[12].

Interval od konce vlny S po peak vlny T je označován jako absolutní refrakterní perioda, během které je srdeční buňka zcela nedráždivá a nereaguje ani na velmi intenzivní vnější podnět. Takto je srdeční myokard chráněn před příliš rychlým opakováním vzruchů, při kterém by hrozilo snížení srdeční čerpací funkce [1],[9].

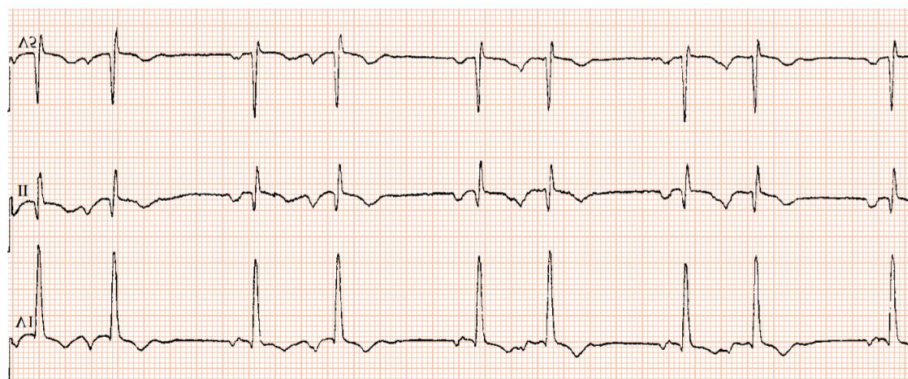
Po vlně T občas následuje plochá vlna U, která nemá úplně jistý původ a její význam není zcela jasný. Nejspíše je projevem repolarizace Purkyňových vláken [3],[8].

1.3 Srdeční extrasystoly

Extrasystola, také označovaná jako ektopický či předčasný stah, je vzruch, který vzniká předčasně v srdečním cyklu mimo SA uzel v ektopickém ložisku v síních, AV junkci nebo v komorách. Vzniká tedy dříve, než by byl očekáván další normálně časovaný srdeční stah a vazebný interval (interval mezi extrasystolou a předcházejícím normálním stahem) je proto kratší než délka cyklu hlavního srdečního rytmu [13].

1.3.1 Síňové extrasystoly

Síňová extrasystola, viditelná na obr. 1.3, je obvykle převedena na komory přes AV uzel, Tarawova raménka a Purkyňova vlákna, QRS komplex je proto štíhlý a jeho tvar není změněn. P vlna odpovídající předčasnému stahu má odlišný tvar než vlna P pocházející z SA uzlu, může dokonce nasedat na předcházející vlnu T a deformovat tak její tvar. Délka intervalu PQ může být zachována nebo zkrácena [13].



Obrázek 1.3 Síňová extrasystola v druhém, čtvrtém, šestém a osmém srdečním úderu (převzato z [13])

1.3.2 Komorové extrasystoly

Příčiny a klinický význam komorových extrasystol

Komorové extrasystoly (KES) jsou nejčastější komorová arytmie v populaci. Vyskytují se u zdravých osob bez prokázaného srdečního onemocnění i u pacientů s různými onemocněními srdce, jsou častější u mužů a jejich prevalence se zvyšuje s věkem. Podle studie [14] z roku 2002 je KES při standardním 12svodovém vyšetření detekováno u 1 % normální zdravé populace, při dvouminutovém EKG vyšetření u 6 % a při 24–48hodinovém Holterovském monitorování až u 40–75 %. V převážné většině případů je KES benigní, závažnost se zvyšuje s přítomností strukturálního onemocnění srdce [15],[16].

K příčinám komorové extrasystolie může patřit akutní infarkt či ischemie myokardu, hypertenze, chlopenní vady, dilatační či hypertrofická kardiomyopatie nebo myokarditida. Mezi mimosrdeční příčiny pak lze zařadit elektrolytovou nerovnováhu (hypokalémie, hypomagnésemie, hyperkalcemie), vliv medikace, alkoholu, kofeinu nebo kouření či vysoké hladiny stresu [11],[17].

Extrasystoly jsou většinou asymptomatické, u některých pacientů se ale mohou příznaky nepříjemně projevit palpací, dušností, závratí nebo bolestí či tlakem na hrudi. U části pacientů může vést četný výskyt KES (pokud KES představuje více než 10 % všech srdečních úderů) až k dilatační kardiomyopatii a tím snížení systolické funkce levé komory srdeční [12],[13],[15].

Projevy v EKG záznamu

Komorové extrasystoly vznikají abnormální elektrickou depolarizací z ektopického ložiska myokardu v oblasti pod větvením Hisova svazku [15]. Vzruch u předčasného komorového stahu není převáděn běžnými drahami pomocí Hisova svazku, Tawarových ramének a Purkyňových vláken, ale pomaleji pracovním myokardem komor [17]. Na EKG záznamu se tedy KES projeví předčasnými širokými QRS komplexy, které trvají více než 120 ms a mají abnormální tvar. QRS komplexu obvykle nepředchází předčasná vlna P, naopak může být ojediněle lokalizována až za ním. QRS komplex je většinou následován úplnou kompenzační pauzou. Vlna T je obvykle vysoká a je spolu s ST úsekem diskordantní vůči orientaci QRS komplexu (jejich výchylky mají opačnou polaritu) [9],[11]. KES se dají popsat a klasifikovat podle několika kritérií.

Klasifikace podle místa vzniku

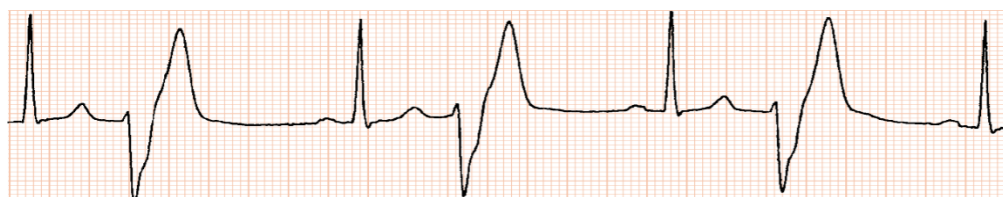
Tvar extrasystoly souvisí s místem vzniku vzruchu. Unifokální KES vznikly v jednom ektopickém ložisku (fokusu), mají proto identický tvar (jsou monomorfní) i stejně dlouhý vazebný interval [18]. Pokud se v myokardu vyskytuje více ektopických fokusů, které generují předčasné impulsy, jsou extrasystoly označovány jako multifokální (multiformní). Kvůli rozdílným průběhům šíření elektrických impulsů v komorách jsou charakterizovány vzájemně odlišnou morfologií (jsou polymorfní) a rozdílnými vazebnými intervaly (viz obr. 1.4) [13],[17].



Obrázek 1.4 Multifokální komorová extrasystola (převzato z [13])

Klasifikace podle frekvence

Bigeminie značí pravidelně se střídající sinusové stahy a komorové extrasystoly, kdy je každý druhý srdeční úder KES (viz obr. 1.5).



Obrázek 1.5 Komorová bigeminie (převzato z [13])

Trigeminie (viz obr. 1.6) je charakterizována jako dva sinusové stahy následované komorovou extrasystolou.



Obrázek 1.6 Komorová trigeminie (převzato z [13])

Kuplet (pár) znamená dvě KES následující v řadě za sebou, jak je vidět na obr. 1.7.

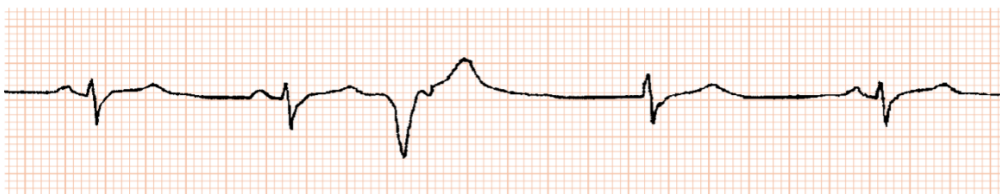


Obrázek 1.7 Kuplet (převzato z [13])

V případě tří KES se mluví o salvě nebo tripletu, dvě a více KES za sebou jsou také často označovány jako nesetrvalá komorová tachykardie. Pokud tachykardie trvá více než 30 sekund nebo vede do 30 sekund k oběhovému kolapsu, je nazývána jako setrvalá komorová tachykardie [13],[18],[19].

Klasifikace podle síňové aktivity

Délka vazebného intervalu stahu následujícího po KES závisí na převodním poměru v AV uzlu a reakci SA uzlu [15]. V případě retrográdního šíření depolarizace z KES AV uzlem z komor na síně, kdy dorazí vzruch do SA uzlu dříve, než se sám stihne depolarizovat, může být v EKG záznamu pozorována invertovaná vlna P nasedající na extrasystolu za QRS komplex (viz obr. 1.8). Dojde k neúplné kompenzační pauze, kdy je součet RR intervalů před a po KES menší než dvojnásobek obvyklého RR intervalu. Tento případ není příliš častý [13],[16].



Obrázek 1.8 Invertovaná vlna P viditelná za třetím QRS komplexem (převzato z [13])

Úplná kompenzační pauza, kdy je součet RR intervalů před a po KES roven dvojnásobku normální délky srdečního cyklu, je pozorovatelná, pokud dojde v AV uzlu ke kolizi retrográdní depolarizační vlny z KES a dopředné vlny generované SA uzlem. Obě vlny se díky refrakterní periodě neprojeví, AV uzel nepřivede vzruch z komor na síně a aktivita síní tak nezávisí na komorách.

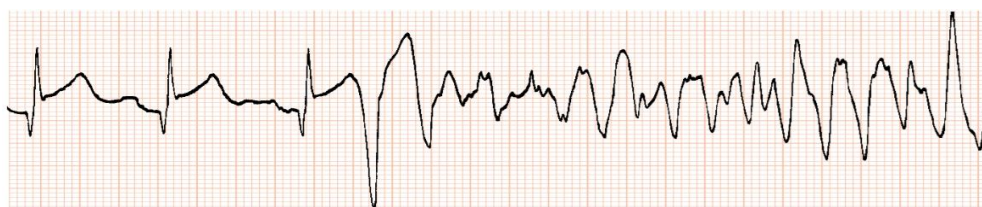
Při interpolovaných KES se může vyskytnout jev označovaný jako skryté retrográdní vedení. Vzruch z komor nezpůsobí retrográdní aktivaci síní, do AV uzlu pronikne pouze částečně. Pokud následující vzruch z SA uzlu zastihne AV uzel v částečné refrakterní fázi, bude převeden pomaleji a PQ interval se tak prodlouží. Kompenzační pauza v tomto případě úplně chybí (viz obr. 1.9) [13],[19].



Obrázek 1.9 Interpolovaná komorová extrasystola s prodlouženým následujícím PQ intervalem, kompenzační pauza chybí (převzato z [13])

Klasifikace podle načasování

Komorové extrasystoly označované jako fenomén „R na T“ vznikají v srdečním cyklu předčasně a mohou tak nasedat na vlnu T předcházejícího srdečního stahu v její vulnerabilní fázi (viz obr. 1.10). V některých případech (při akutním infarktu myokardu, extrémní hypokalémii nebo při prodlouženém QT intervalu) mohou vyvolat komorovou tachykardii nebo dokonce fibrilaci komor [19].



Obrázek 1.10 Fenomén „R na T“ vyvolávající fibrilaci komor (převzato z [13])

End-diastolická KES vzniká na konci diastoly bezprostředně po vlně P vyvolané normální aktivací SA uzlu těsně před očekávaným sinusovým QRS komplexem. Komory jsou aktivované z KES, síně z SA uzlu. Po ní následuje úplná kompenzační pauza (viz obr. 1.11) [13],[17].



Obrázek 1.11 End-diastolická komorová extrasystola (převzato z [13])

Interpolovaná extrasystola (viz obrázek 1.9) se vyskytuje zcela výjimečně při obvykle pomalejším pravidelném sinusovém rytmu, je vmezeřená mezi dva normální stahy

a neprodukuje žádnou kompenzační pauzu. PQ interval je po interpolované KES mírně prodloužený [13],[16].

Klasifikace dle Lowna

KES jsou nejčastěji klasifikovány podle Lownova rozdělení do pěti tříd podle závažnosti (viz tabulka 1.1). Se zvyšujícím se počtem a komplexností zjištěných KES roste z klinického hlediska význam a závažnost. Za komplexní formy se pokládají stupně 3 až 5. Tento způsob rozdělení nemusí být vždy vhodný, protože byl vytvořen pro pacienty s akutním koronárním syndromem [19]. Za ojedinělé KES se pokládá výskyt méně než 5 KES za hodinu, za čtenější 5-39 KES/hod, za časté 40-140 KES/hod a za velmi časté více než 140 KES/hod. Jednotlivé izolované KES se vyskytují u zdravých osob, četný výskyt více KES po sobě může vypovídat o akutním infarktu, ischemii, srdečních vadách či poruchách chlopní [12],[16],[18].

Tabulka 1.1 Lownova klasifikace komorových extrasystol

Stupeň	Typ KES
0	žádné KES
1a	izolované, sporadické, unifokální KES, < 1/min a < 30/hod
1b	izolované, sporadické, unifokální KES, > 1/min, ale < 30/hod
2	izolované, časté, unifokální KES, > 30/hod
3a	izolované, multifokální KES
3b	bigeminicky vázané uni- či multifokální KES
4a	KES v párech (kuplety)
4b	KES v salvách, tj. 3 a více následných KES
5	časné KES (fenomén „R na T“)

Při výskytu ektopického ložiska ve svalovině pravé komory připomíná tvar KES blokádu levého Tawarova raménka. KES vznikající v levé komoře se naopak podobají blokáde pravého Tawarova raménka. Toto by mělo být bráno v úvahu při návrhu a testování algoritmu pro detekci KES [16],[17].

1.4 Raménkové blokády

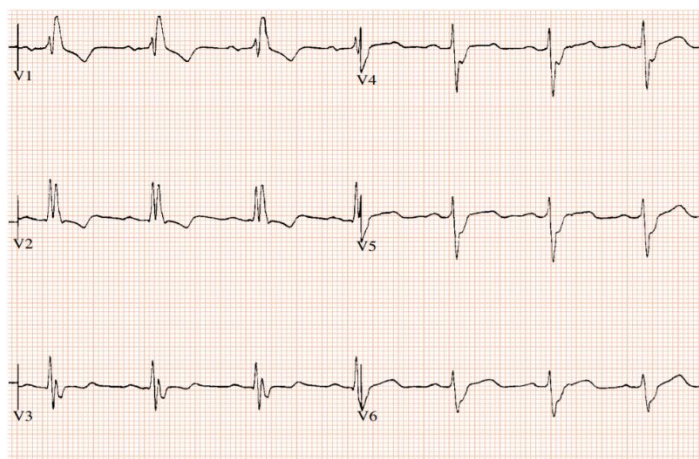
V praxi se blokády obou Tawarových ramének vyskytují poměrně běžně. Normálně se vzruch do komor šíří přes AV uzel a Hisův svazek, který se rozděluje na levé a pravé Tawarovo raménko, což umožňuje velmi rychlou a téměř současnou depolarizaci obou komor. Porucha vedení vzruchu přes některé z ramének ústí v opožděnou a neuspořádanou depolarizaci myokardu komor, jelikož se vzruch přes nepoškozené raménko a Purkyňova vlákna šíří na myokard odpovídající komory, a až poté pasivně

přechází na svalovinu druhé komory, která měla být aktivně depolarizována poškozeným raménkem. To se na EKG záznamu projevuje prodlouženým QRS komplexem s abnormálním tvarem [12]. O zpoždění komorového vedení se jedná v případě, kdy depolarizace komor trvá déle než 100 ms. Při částečné (inkompletní) blokádě je délka QRS komplexu přibližně 100 až 120 ms, u úplné (kompletní) raménkové blokády dochází k rozšíření QRS komplexu na více než 120 ms. Platí, že šířka QRS komplexu je úměrná rozsahu blokády [20]. Samotné raménkové blokády jsou hemodynamicky nevýznamné, mohou ale signalizovat poškození srdce [9],[13].

1.4.1 Blokáda pravého raménka Tawarova

Při blokádě pravého raménka Tawarova (BPRT) je opožděna depolarizace pravé komory, zatímco aktivace mezikomorového septa a levé komory probíhá fyziologicky. Začáteční část QRS komplexu má proto normální tvar [9]. Kvůli opožděné aktivaci pravé komory je QRS komplex rozšířený na více než 120 ms. Protože je pravá komora depolarizována později než levá, je v některých svodech přítomna sekundární R vlna, což vyústí v QRS komplex tvaru M, který může vzhledem připomínat králičí uši (viz obr. 1.12) V několika svodech je pozorovatelný široký kmit S nebo diskordantní úsek ST a vlna T. Následkem částečné blokády pravého raménka je podobný EKG obraz, ale délka QRS komplexu je 100 až 110 ms [12].

BPRT je poměrně častá u lidí s jinak zdravým srdcem, vyskytuje se u pacientů s vrozenou srdeční vadou, může být izolovaným vrozeným poškozením převodního systému. Mezi vážnější příčiny patří hypertrofie pravé komory nebo její přetížení, například při chronickém obstrukčním onemocnění plic nebo při poškození myokardu. BPRT je běžná u onemocnění, které poškozují převodní systém. Podle současných poznatků BPRT nejspíše není negativně prognosticky významná, nicméně čtyřnásobně zvyšuje riziko rozvoje AV blokády [13].

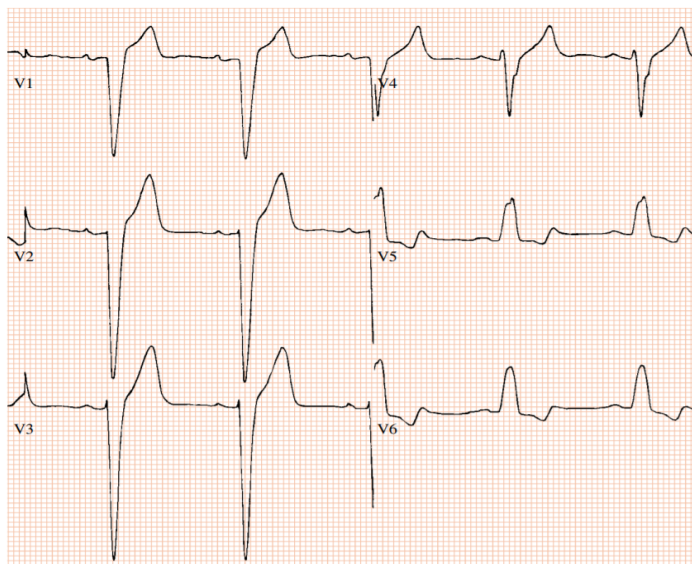


Obrázek 1.12 Blokáda pravého Tawarova raménka (převzato z [13])

1.4.2 Blokáda levého raménka Tawarova

U blokády levého raménka Tawarova (BLRT) je opožděna aktivace levé komory. Depolarizace mezikomorového septa probíhá zprava doleva, tedy v opačném směru než normálně. Pravá komora je depolarizována jako první a elektrická aktivace se poté postupně pomaleji šíří přes svalovinu levé komory. V EKG záznamu je proto první malý negativní kmit q v některých svodech nahrazen větším pozitivním kmitem R. Kvůli zpožděné depolarizaci levé komory je QRS komplex rozšířený (≥ 120 ms, viz obr. 1.13) a kmit R je v některých svodech široký a rozštěpený, někdy může být taky projevem invertovaná vlna T. Částečná blokáda levého raménka má podobný EKG obraz, ale délka QRS komplexu je 100–110 ms.

BLRT je na rozdíl od BPRT s větší pravděpodobností spojeno se strukturálním onemocněním srdce. Levé Tawarovo raménko je totiž ve srovnání s pravým mohutnější, a pro vznik převodní blokády je tak potřeba poškození výrazného množství tkáně. Příčinou BLRT může být hypertrofie levé komory nebo poškození myokardu kvůli ischemické chorobě srdeční (ICHS) či kardiomyopatii, často se vyskytuje u onemocnění převodního systému. BLRT 18násobně zvyšuje riziko rozvoje AV blokády a srdečního selhání, je spojená se zvýšeným rizikem úmrtí, hlavně náhlé smrti při ICHS [9],[13].



Obrázek 1.13 Blokáda levého Tawarova raménka (převzato z [13])

2. METODY DETEKCE KOMOROVÝCH EXTRASYSTOL

Detekce srdečních arytmií představuje v současné době jeden z velmi důležitých úkolů. Přesná a včasná diagnostika problému je klíčová vzhledem k riziku rozvinutí vážnějších srdečních onemocnění a k případné další léčbě. Správné vyhodnocení EKG záznamu závisí na odborných znalostech a zkušenostech lékaře, může být časově náročné.

V poslední době bylo z těchto důvodů navrženo mnoho metod automatické detekce komorových extrasystol, které se dají rozdělit do dvou hlavních skupin. Podstatou algoritmů patřících do první skupiny je výběr a výpočet relevantních příznaků, které souvisí s typickými projevy KES v EKG záznamu (viz kapitola 1.3.2). Pomocí nich jsou poté jednotlivé srdeční úderů různými metodami klasifikovány. Do druhé skupiny se řadí algoritmy využívající hlubokého učení, ve kterých předchozí selekce příznaků není potřebná, jelikož si síť významné příznaky vybere z EKG signálu sama. Do první skupiny jsou zařazeny metody, které ke klasifikaci také využívají hluboké učení, nicméně v nich vstupem do sítě nebyl samotný EKG záznam, ale z něj vypočtené příznaky. Ve většině metod se kromě KES detekují i další typy arytmií a poruch srdečního převodu.

Cílem této části je představení velkého množství různorodých jednoduchých i složitějších přístupů a metod klasifikace srdečního rytmu k získání lepší představy o jejich principu, výhodách a nevýhodách. Získané znalosti a poznatky jsou využity při vytváření vlastní metody klasifikace v kapitole 4. V kapitole 5.3 je popsán způsob rozdělení dat, průběh trénování i dosažené výsledky většiny metod zmíněných v této kapitole.

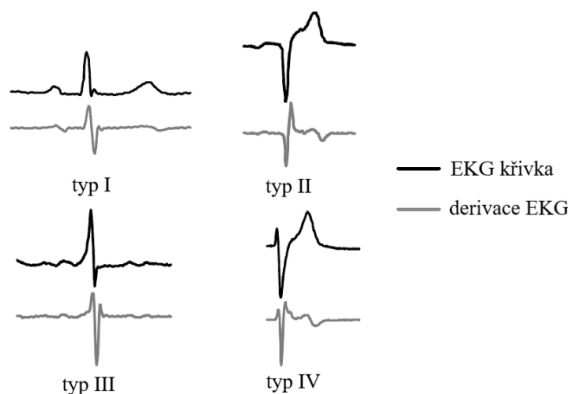
2.1 Metody založené na výběru relevantních příznaků

2.1.1 Metody založené na aplikaci rozhodovacích pravidel

V *první jednoduché metodě* popsané v článku [21] je detekce KES založena na výpočtu tří příznaků a následné klasifikaci pomocí rozhodovacích pravidel. Prvním z příznaků je RR-interval, tedy délka mezi dvěma po sobě následujícími R vlnami. Další příznak je QRS-vzor, který je z EKG křivky spočítán jako její derivace (viz obrázek 2.1). Na základě svého tvaru je pak QRS komplex zařazen do jednoho ze čtyř typů vzorů. Posledním příznakem je QRS-šířka. Detekce začátku a konce QRS komplexu závisí na určeném typu QRS-vzoru a sklonu derivovaného signálu.

Jednotlivé srdeční úderů jsou klasifikovány pomocí rozhodovacích pravidel. Pokud je aktuální RR-interval kratší než $0,94 \cdot ARR_{n-1}$, kde ARR_{n-1} je průměr osmi předchozích normálních RR-intervalů, a zároveň je QRS-šířka delší než $1,32 \cdot AW_{n-1}$, kde AW_{n-1} je průměr osmi předchozích normálních QRS-šířek, je daný srdeční úder klasifikován jako KES. Pokud platí pouze první podmínka, je porovnán aktuální QRS-vzor se vzory osmi předchozích úderů neklasifikovaných jako KES. Pokud je typ vzoru předchozích osmi normálních úderů stejný a vzor aktuálně zkoumaného srdečního úderu

se liší, je srdeční úder klasifikován jako KES. Ve všech ostatních případech je aktuální úder klasifikován jako normální. Mezi výhody metody patří její jednoduchá implementace a možná aplikace v reálném čase, nevýhodou pak je velké množství špatně klasifikovaných úderů.



Obrázek 2.1 QRS vzory (převzato a upraveno z [21])

Následující jednoduchá metoda [22] detekce extrasystol pro nositelná EKG zařízení ke klasifikaci srdečního rytmu také využívá rozhodovacích pravidel. EKG záznam je nejprve filtrován pásmovou propustí s propustným pásmem 0,1–45 Hz, následně jsou detekovány R peaky a EKG záznam je rozdělen na segmenty dlouhé 10 s.

Samotný detekční algoritmus sestává ze tří fází. Na základě R peaků jsou pro každý srdeční úder odvozeny 4 parametry: délka RR intervalů, výška, šířka a šablona QRS komplexu. Šířka QRS komplexů je měřena ve dvou pětinach výšky nejnižšího QRS komplexu, šablona QRS komplexu je pak vybrána jako úsek o délce 0,2 s se středem v R peaku. Druhou fází algoritmu je hrubá detekce, kdy jsou jednotlivé srdeční údery na základě rozhodovacích pravidel, které využívají předem zmíněných parametrů, klasifikovány jako normální rytmus či komorové nebo supraventrikulární extrasystoly (SVES). Poslední fází algoritmu je jemná detekce, jejíž úkolem je nalezení špatně klasifikovaných úderů. Výhodou algoritmu je nízká výpočetní náročnost, jednoduchá extrakce příznaků a možnost použití v reálném čase, nevýhodou pak nepřesná klasifikace při více druzích arytmiích, které jednoduché příznaky nedokážou přesně postihnout.

2.1.2 Metody využívající časových, morfologických či statistických příznaků

Metoda prezentovaná v článku [23] ke klasifikaci různých srdečních arytmií používá tři typy příznaků. Po odstranění kolísání nulové izolinie a detekci R vln jsou segmentovány jednotlivé srdeční údery tak, že jsou vybrány vzorky z intervalu 250 ms před R peakem a 400 ms po něm, z každého úderu je tedy vybrán úsek dlouhý 0,65 s.

Prvním typem příznaků jsou RR intervaly, jmenovitě pre-RR, který je definován jako časová vzdálenost mezi aktuálním a předchozím R peakem, a post-RR neboli vzdálenost mezi současným a následujícím R peakem. Dalším typem příznaků jsou centrální momenty vyššího řádu, které obecně charakterizují hustotu rozdělení pravděpodobnosti

náhodné veličiny. Z nich jsou vybrány 3: koeficient šikmosti, koeficient špičatosti a pátý centrální moment. Poslední typ příznaků jsou parametry 2D GMM¹, které jsou pro každý segment optimalizovány pomocí EM² algoritmu. Mezi parametry patří váha (prakticky amplituda), střední hodnota a rozptyl pro dvě Gaussovy křivky. Každý srdeční úder je na základě příznakového vektoru následně klasifikován pomocí algoritmu rozhodovacích stromů.

Následující přístup [24] využívá ke klasifikaci pěti různých typů srdečního rytmu morfologické i dynamické příznaky. Po předzpracování signálu a nalezení R peaků je záznam rozdělen na úseky dlouhé 300 vzorků skládající se ze 100 vzorků před R vlnou a 200 vzorků za ní. Na každý segment je následně aplikována čtyřstupňová diskrétní vlnková transformace (DWT³), která signál dokáže charakterizovat v časové i frekvenční oblasti, a je tak efektivní při analýze nestacionárních signálů. Detailní koeficienty na 3. a 4. úrovni rozkladu a aproximační koeficienty 4. úrovně rozkladu jsou použity jako jedna část morfologických příznaků. Z každého segmentu je pomocí analýzy nezávislých komponent (ICA⁴) odvozeno 14 koeficientů reprezentujících další morfologické příznaky. Pro každý segment reprezentující jeden srdeční úder je tak získáno dohromady 128 morfologických příznaků. Pro snížení jejich počtu je použita analýza hlavních komponent (PCA⁵), jejíž výstupem je výsledných 18 hlavních komponent. Dále jsou z RR intervalů odvozeny 4 příznaky charakterizující dynamické vlastnosti každého úderu srdce. První je délka předchozího RR intervalu, druhý délka následujícího RR intervalu, dalším je místní RR interval spočítaný jako průměr RR intervalů v předchozích 10 vteřinách signálu, a poslední dynamický příznak je průměrný RR interval neboli průměr RR intervalů v předchozích pěti minutách signálu.

Každý segment reprezentovaný 18 morfologickými a 4 dynamickými příznaky je následně klasifikován pomocí metody podpůrných vektorů (SVM⁶) s radiální bázovou jádrovou funkcí (RBF), jejíž výstupem je pravděpodobnost příslušnosti ke každé třídě. Jelikož jsou použity záznamy ze dvou svodů, je popsán postup použit nezávisle pro oba signály. Výsledná klasifikace segmentu je pak určena spojením výstupů z obou SVM klasifikátorů.

2.1.3 Metody využívající analýzu hlavních komponent

Jedna z metod využívá analýzu hlavních komponent a metodu podpůrných vektorů [25]. Jednotlivé srdeční údery jsou klasifikovány do pěti kategorií: normální rytmus, blokáda levého a pravého Tawarova raménka, komorová a supraventrikulární extrasystola. Po předzpracování signálu a detekci kmitů R je EKG záznam rozdělen do úseků o délce 200

¹ z anglického Gaussian Mixture Model

² z anglického Expectation-Maximization (očekávaná střední hodnota-maximalizace)

³ z anglického Discrete Wavelet Transform

⁴ z anglického Independent Component Analysis

⁵ z anglického Principal Component Analysis

⁶ z anglického Support Vector Machine

vzorků tak, že se vždy vezme 99 vzorků před aktuálním kmitem R a 100 vzorků za ním. Z každého segmentu je díky PCA vybráno pouze 12 nejvýznamnějších hlavních komponent, které jsou následně použity pro klasifikaci. Příznaky jsou do jednotlivých tříd rozřazeny pomocí nelineární SVM, kde je jako jádrová funkce použita RBF.

Další zajímavá metoda využívá pro detekci KES směr největšího rozptylu v datech vypočítaný z PCA [26]. V EKG záznamu jsou nejprve detekovány kmity R a pomocí nich jsou vybrány úseky signálu tak, že každý úsek obsahuje 50 vzorků před kmitem R a 99 vzorků za ním. Při segmentaci tak vzniknou úseky dlouhé 150 vzorků, které jsou následně normalizovány, aby měly nulovou střední hodnotu a směrodatnou odchylku rovnou jedné.

Z prvních dvou minut každého signálu je vybráno 10 úseků normálního rytmu, ze kterých je vytvořena matice M o rozměrech 150×10 . Na matici M je poté aplikována PCA, díky které dostaneme vlastní vektor definující směr největšího rozptylu v datech. Dále postupně vždy jeden ze všech zbývajících segmentů EKG signálu nahradí jeden sloupec (neboli jeden úsek normálního rytmu) v matici M a znovu je spočítán směr největšího rozptylu v datech. Pokud úhel mezi oběma směry překročí stanovený práh, je daný segment klasifikován jako KES. Naopak, pokud je úhel menší než další práh, nahradí uvažovaný segment jeden úsek v matici M . Současně jsou aktualizovány i oba prahy, metoda je tak schopná adaptace na změny tepové frekvence a morfologie QRS komplexů.

2.1.4 Metody založené na vyhledávání vzorů

Podstatou *jedné vybrané metody* [27] je klasifikace srdečních cyklů na základě vyhledávání podobnosti mezi jejich časovými a morfologickými příznaky. Po odstranění kolísání nulové izolinie, síťového brumu a myopotenciálů jsou detekovány R peaky. EKG signál je segmentován tak, že se každý úsek skládá ze 160 vzorků před R peakem a 150 vzorků po něm. Z každého segmentu jsou následně získány 4 parametry. Prvním je délka RR intervalu, druhým doba trvání QRS komplexu, poslední dva parametry jsou amplitudy vln R a T. Po jejich výpočtu je na základě těchto parametrů ze všech úseků vybráno 5 segmentů normálního rytmu, které slouží jako vzory pro detekci KES. Následně je mezi všemi vzory vypočítána normalizovaná vzájemná korelace

$$NCC = \frac{\sum_{i=0}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=0}^N (x_i - \bar{x})^2} \sqrt{\sum_{i=0}^N (y_i - \bar{y})^2}} \quad (2.1)$$

kde x a y jsou segmenty EKG signálu, \bar{x} je průměrná hodnota segmentu x , \bar{y} je průměrná hodnota segmentu y a N je počet vzorků v segmentu. Ze vzájemných korelací vzorů je spočítáno průměrné maximum a minimum korelačního koeficientu. Pro každý další EKG úsek je poté spočítána normalizovaná vzájemná korelace s každým ze vzorových segmentů. Pokud je maximální hodnota korelace vyšší než 0,8, je uvažovaný

segment klasifikován jako normální. V opačném případě jsou porovnány dříve vypočtené parametry daného segmentu a vzoru. Pokud je RR interval uvažovaného segmentu delší než 1,2 s, doba trvání QRS komplexu vyšší než 0,12 s, amplituda T vlny nižší než 0 mV nebo minimální hodnota korelace nižší než průměrné minimum korelačního koeficientu, je segment klasifikován jako KES.

Další metoda využívající vyhledávání vzorů k detekci KES byla popsána v článku [28]. Ke trénování i testování byly použity záznamy dlouhé 30 minut, prvních 5 minut slouží k vytvoření individualizovaného vzoru, detekce KES je pak provedena na zbývajících 25 minutách záznamu. V prvních pěti minutách záznamu jsou spočítány RR intervaly a pro každý z nich je vytvořena RR sekvence tak, že se vezme aktuální a devět následujících RR intervalů. Vznikne tak $N-9$ sekvencí o délce 10 hodnot, kde N je celkový počet RR intervalů v prvních 5 minutách záznamu. Ze sekvence s nejmenší směrodatnou odchylkou je poté vypočítána průměrná délka RR intervalů \overline{RR} . Úseky EKG odpovídající hodnotám RR intervalů z této vybrané sekvence jsou převzorkovány na délku \overline{RR} . Z těchto 10 stejně dlouhých segmentů pak je vytvořen průměrný segment \bar{x} . Ze dvou identických segmentů \bar{x} spojených s překryvem jednoho vzorku je vybrán úsek o délce \overline{RR} se středem v překryvu obou segmentů. Tento úsek představuje individualizovaný vzor pro daný záznam.

Následně je spočítána vzájemná korelační funkce mezi zbývajících 25 minutami signálu a vytvořeným vzorem, která je poté normalizována. Pomocí druhé derivace této funkce jsou nalezeny pozice R peaků, které slouží k segmentaci záznamu na jednotlivé srdeční úder. Z EKG signálu jsou vybrány úseky se středem v R peaku o délce M

$$M = 1 + 2 \cdot \lfloor 0,1 \cdot \overline{RR} \rfloor, \quad (2.2)$$

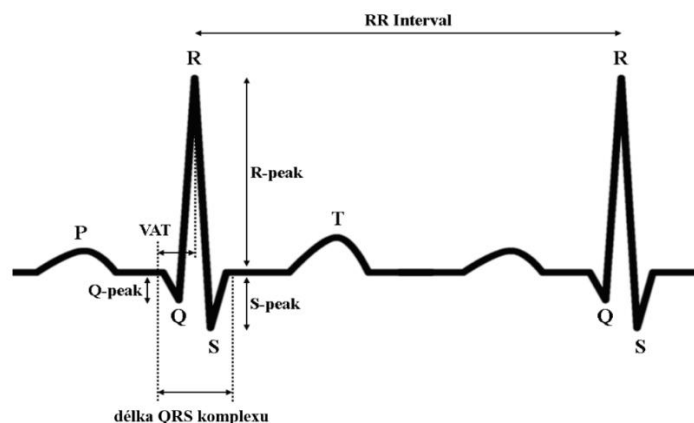
kde $\lfloor \cdot \rfloor$ reprezentuje operátor *floor*, který číslo v argumentu zaokrouhlí na nejbližší celé číslo vždy směrem k nule. Na stejnou délku je oříznut i vytvořený vzor. Nakonec je spočítán normalizovaný korelační koeficient pro každý úsek a oříznutý vzor. Pokud je korelace nižší než stanovený práh, je daný úsek označen jako KES, v opačném případě jako normální.

2.1.5 Metody využívající hluboké neuronové sítě

Základem *metody* publikované v článku [29] je výpočet několika parametrů a klasifikace pomocí hluboké neuronové sítě (DNN⁷). Na EKG signál je nejprve aplikován Pan-Tompkins algoritmus, který sestává ze čtyř kroků: filtrace pásmovou propustí pro potlačení kolísání nulové izoliny, myopotenciálů i síťového brumu, dále pětibodová derivace a umocnění získaného signálu pro snazší detekci R vln. Posledním krokem je integrace v posuvném okně. K algoritmu je v článku jako další krok přidána detekce Q a S peaků z integrovaného signálu. Tímto postupem je pro každý srdeční úder extrahováno 6 příznaků: délka RR intervalu, délka QRS komplexu, doba aktivace komor

⁷ z anglického Deep Neural Network

(VAT⁸), amplituda Q peaku, R peaku a S peaku. Všechny příznaky jsou vyznačeny na obrázku 2.2.



Obrázek 2.2 Příznaky získané z EKG signálu (převzato a upraveno z [29])

Tyto příznaky po min-max normalizaci, která transformuje rozsah hodnot proměnných od 0 do 1, vstupují do DNN, která na jejich základě provede pro každý srdeční cyklus klasifikaci srdečního rytmu. DNN se skládá ze vstupní vrstvy, šesti skrytých vrstev s postupně 10, 30, 50, 100, 40 a 20 neurony a z výstupní vrstvy, která rytmus klasifikuje jako normální nebo KES. V DNN byla použita L2 regularizace, softsign aktivační funkce, optimalizační algoritmus Adam⁹, Xavierova inicializační metoda vah a postupně se snižující krok učení.

Další metoda popsaná v článku [30] ke klasifikaci srdečních arytmií také využívá DNN. Signál je po předzpracování nasegmentován na úseky začínající P vlnou a končící T vlnou, jednotlivé segmenty jsou podvzorkovány na délku 50 vzorků. Pro každý segment reprezentující jeden srdeční úder jsou dále spočítány 4 parametry: pre-RR interval definovaný jako délka RR intervalu mezi předchozím a současným R peakem, post-RR interval neboli délka RR intervalu mezi současným a následujícím R peakem, lokální průměrný RR interval, tedy průměr délek RR intervalů v předcházejících 10 s, a globální průměrný RR interval definovaný jako průměr délek RR intervalů v předcházejících pěti minutách. Tyto parametry spolu s 50 vzorky pro každý srdeční cyklus jsou vstupem do DNN. DNN představená v tomto článku je složena ze vstupní vrstvy, sedmi skrytých vrstev s ReLU¹⁰ aktivační funkcí s postupně 5, 10, 30, 50, 30, 10 a 5 neurony a z výstupní vrstvy se softmax aktivační funkcí, která úseky klasifikuje jako normální nebo arytmiické. Jako kriteriální funkce byla zvolena křížová entropie.

2.1.6 Metody založené na výpočtu korelace

Metoda prezentovaná v článku [31] je založena na analýze cyklostacionárních signálů,

⁸ z anglického Ventricular Activation Time

⁹ z anglického Adaptive moment estimation

¹⁰ z anglického Rectified Linear Unit

kteřá předpokládá, že EKG signál není stacionární. Tato analýza je v signálu schopná odhalit skryté periodicity, navíc na ni nemá vliv šum, který nevykazuje cyklickou povahu.

EKG záznam není předzpracován, je pouze rozdělen na segmenty dlouhé 300 vzorků, přesněji 110 vzorků před R peakem a 189 vzorků po něm. Pro každý segment je vypočítána spektrální korelace, vznikne tak matice spektrálních korelačních koeficientů, ve které řádky reprezentují frekvenci a sloupce cyklickou frekvenci. Nenulové hodnoty v matici značí to, že signál vykazuje periodicitu. Jelikož je většina hodnot v matici přibližně nulová, je prahována pomocí Otsu metody tak, že je matice oříznuta a tím jsou zachovány pouze spektrální koeficienty obsahující nejvíce informací. Tyto koeficienty pak tvoří vstupní příznakový vektor. Klasifikace je provedena pomocí nelineární SVM s lineární jádrovou funkcí.

Následující metoda založená na výpočtu vzájemné korelace byla publikována v článku [32]. Po předzpracování signálu a detekci R peaků jsou vysegmentovány jednotlivé srdeční úderý. Z těchto úseků jsou vybrány pouze ty, ve kterých nejsou prudké změny v izolínii. Dále je spočítána vzájemná korelace mezi všemi segmenty, a podle ní jsou segmenty rozděleny do shluků s tím, že každý může náležet pouze do jednoho shluku. Klasifikace je provedena aplikací dvou pravidel. První je založeno na časových vlastnostech signálu, druhé na rozdílném tvaru pravidelných a nepravidelných srdečních úderů. Mezi výhody této metody patří fakt, že umožňuje individuální analýzu každého záznamu a nepotřebuje trénovací sadu dat. Přesnost klasifikace tak není snížena kvůli rozdílnému tvaru QRS komplexů stejného rytmu mezi trénovacími daty či výběrem trénovacích dat. Metoda je zároveň velmi jednoduchá a nevyžaduje ke klasifikaci velké množství příznaků či složité algoritmy. Nevýhodou pak je, že nepokřývá případ, kdy je v signálu více KES než normálních úderů.

2.2 Metody využívající hluboké učení

Metody hlubokého učení jsou v současné době používány nejčastěji, využívají neuronové sítě s velkým množstvím skrytých vrstev, díky kterým si jsou schopny v průběhu trénování samy extrahovat relevantní příznaky. Výhodou těchto metod je tedy to, že při jejich použití není třeba předem počítat z dat příznaky. Navíc jsou schopny řešit i komplexní problémy. Mezi nevýhody se pak dá řadit velká výpočetní náročnost, nutnost použití velkého množství dat k natrénování, náročná optimalizace hyperparametrů, navíc u nich poměrně lehce dojde k přeučení. U příznaků extrahovaných sítí taky chybí přímý či jednoznačný důkaz o jejich medicínském významu. Ke klasifikaci srdečních arytmií se nejčastěji využívají konvoluční nebo rekurentní neuronové sítě.

2.2.1 Metody využívající autoenkodéry

Jedna z metod [33] využívá k detekci KES konvoluční autoenkodér (CAE¹¹) a náhodný

¹¹ z anglického Convolutional AutoEncoder

les. EKG signály jsou nejprve rozděleny na segmenty o délce 250 vzorků, kde 89 vzorků je před R peakem a 160 vzorků za ním. CAE je použit k extrakci příznaků, na jejichž základě je následně segment klasifikován pomocí metody náhodného lesa jako normální nebo KES.

Po natrénování CAE byly ve výsledném algoritmu použity pouze první dvě konvoluční vrstvy s lineárními aktivacemi. První vrstva obsahuje 25 konvolučních jader o délce 20 vzorků s délkou kroku 10, druhá jedno konvoluční jádro o délce 25 vzorků. Výstupem z CAE je tedy vektor o délce 25 vzorků, které reprezentují jednotlivé příznaky a každý je lineární kombinací výstupů z první vrstvy. Segment je klasifikován pomocí metody náhodného lesa, který se skládal z 10 rozhodovacích stromů s chybovou funkcí Gini impurity, zároveň byl pro redukci přeučení použit bagging¹².

Tato další metoda [34] využívá ke klasifikaci pěti typů srdečního rytmu konvoluční autoenkodér, který zároveň slouží k odstranění šumu (CDAE¹³). EKG záznamy jsou nejprve rozděleny na jednotlivé segmenty se středem v R peaku a počtem vzorků odpovídajícím délce předchozího RR intervalu. Jelikož jsou takto vybrané úseky různě dlouhé, jsou převzorkovány na jednotnou délku 160 vzorků a rozsah hodnot je následně normalizován od 0 do 1. Pro následnou klasifikaci jsou použity signály ze dvou svodů. CDAE složený ze sedmi vrstev ze vstupních dat extrahuje pouze důležité a robustní příznaky. První vrstva CDAE signál poškodí šumem s normálním rozložením hodnot. Další tři dvojice vrstev jsou vždy konvoluční s ReLU aktivační funkcí a max-pooling vrstva. Za třetí max-pooling vrstvou jsou přidány dvě plně propojené vrstvy se softmax aktivační funkcí, které slouží ke klasifikaci srdečního rytmu.

Následující metoda popsaná v článku [35] provádí klasifikaci šesti typů rytmu (normální, BPRT, BLRT, KES, SVES a rytmus udávaný kardiostimulátorem) za pomoci čtyř úrovní autoenkodérů poskládaných za sebe (SSAEs¹⁴). Každá úroveň se skládá ze tří vrstev. První je tzv. viditelná vrstva sloužící jako enkodér. Výstupem druhé (tzv. skryté) vrstvy jsou extrahované příznaky dané úrovně. Třetí vrstva (dekodér) je použita pouze při trénování sítě, slouží k rekonstrukci původního signálu vstupujícího do autoenkodéru na dané úrovni z příznaků ze skryté vrstvy. Architektura sítě je taková, že skrytá vrstva prvního autoenkodéru je zároveň viditelná vrstva autoenkodéru v druhé úrovni, a skrytá vrstva druhého autoenkodéru je stejně tak viditelná vrstva třetí úrovně. Trénování SSAEs probíhá postupně. Nejprve je natrénován autoenkodér na nejnižší úrovni. Příznaky ze skryté vrstvy první úrovně jsou použity jako viditelná vrstva druhé úrovně, a na jejich základě je natrénován druhý autoenkodér. Stejným postupem jsou natrénovány všechny čtyři autoenkodéry. Výstupem z nich jsou příznaky ze skryté vrstvy na čtvrté úrovni, na které je nakonec ke klasifikaci srdečního rytmu aplikována softmax regresní funkce.

¹² z anglického bootstrap aggregating

¹³ z anglického Convolutional Denoising AutoEncoder

¹⁴ z anglického Stacked Sparse AutoEncoders

Nevýhodou metody je velký počet vah sítě, který má za následek vyšší výpočetní náročnost a množství dat nutných k trénování, navíc je zvýšeno riziko přeučení sítě.

2.2.2 Metody využívající konvoluční neuronové sítě

Následující metody ke klasifikaci srdečního rytmu používají konvoluční neuronové sítě (CNN¹⁵), liší se pouze v architektuře navržených sítí a nastavených parametrech.

Metoda [36] využívá ke klasifikaci pěti typů srdečního rytmu hluboké konvoluční neuronové sítě. Po odstranění šumu jsou signály segmentovány na úseky o délce 260 vzorků se středem v R peaku a normalizovány za použití Z-skóre. Každý segment je použit jako vstup do CNN, která sestává z 9 vrstev: 1., 3. a 5. vrstva jsou konvoluční, 2., 4. a 6. max-poolingové a poslední tři jsou plně propojené. Jako aktivační funkce pro konvoluční a první dvě plně propojené vrstvy byla použita leaky ReLU, pro výstupní vrstvu sítě pak softmax aktivace. CNN model byl natrénován algoritmem zpětného šíření chyby. Mezi výhody metody patří robustnost vůči šumu a špatné kvalitě signálu, navíc je metoda plně automatická, jelikož není nutné ze signálu předem extrahovat příznaky. Nevýhodou pak je velká výpočetní náročnost a značné množství vzorků nutné k natrénování sítě tak, aby byla schopná v signálu spolehlivě rozpoznávat různé vzory.

CNN byla použita i pro klasifikaci KES, BPRT a normálního srdečního rytmu v článku [37]. Po nalezení R peaků jsou signály segmentovány na úseky dlouhé 300 vzorků tak, že R vlna je 100. vzorkem úseku. Takto segmentovaný signál je následně použit jako vstup do devítivrstvé CNN, kde 1., 3., 5. a 7. vrstvy jsou konvoluční s ReLU aktivační funkcí, vrstvy po nich následující jsou mean-poolingové a poslední vrstva, jejíž výstupem je klasifikace segmentů do jednotlivých tříd, je plně propojená se softmax aktivační funkcí.

V článku [38] bylo CNN využito ke klasifikaci signálů na normální, BPRT, BLRT, KES a SVES. EKG záznam je nejdříve předzpracován a jsou v něm detekovány kmity R. Následně je signál segmentován na úseky dlouhé 201 vzorků s R peakem uprostřed, tyto segmenty jsou poté převzorkovány na délku 130 vzorků. Takto upravené úseky vstupují do CNN s pěti vrstvami. První a třetí vrstva jsou konvoluční, druhá a čtvrtá poolingové, poslední vrstva je plně propojená.

¹⁵ z anglického Convolutional Neural Network

3. HLUBOKÉ UČENÍ

V této kapitole jsou stručně představeny konvoluční neuronové sítě a sítě typu long-short term memory (LSTM¹⁶). Cílem je tak dodat teoretický základ k následující kapitole věnující se praktické části práce, kde je využita kombinace obou těchto sítí.

Síť typu CNN byla zvolena na základě předchozí rešerše pro svou schopnost řešit komplexní problémy a s vysokou úspěšností klasifikovat nová data. U této sítě také odpadá potřeba z dat předem počítat relevantní příznaky, protože si je schopná v průběhu trénování extrahovat sama. Další velkou výhodou je rychlejší trénování ve srovnání s klasickými DNN o stejném počtu vrstev z důvodu nižšího množství trénovatelných parametrů. Ke klasifikaci byla využita síť LSTM, která je jedním z typů rekurentních neuronových sítí. Tato síť má zpětnovazebné konekce, je vhodná ke zpracování sekvencí a signálů. Mezi nevýhody obou sítí se dá oproti klasickým metodám řadit velká výpočetní náročnost, nutnost použití značného množství dat k natrénování a obecně velká závislost na správném předzpracování a rozložení dat. U neuronových sítí je také poměrně náročná optimalizace hyperparametrů, navíc jsou sítě náchylné k přeučení. Všechny zmíněné nedostatky vyžadují při tvorbě sítě velkou pozornost.

3.1 Konvoluční neuronové sítě

Konvoluční neuronové sítě jsou typem hlubokých dopředných neuronových sítí. Základem CNN, které jsou v poslední době využívány velmi často k řešení mnoha typů úloh, jsou konvoluční vrstvy, podle kterých neuronová síť (NN¹⁷) také dostala své jméno. CNN oproti klasickým plně propojeným neuronovým sítím používají méně vah a jsou méně výpočetně náročné. Umí řešit komplexní úlohy, díky své schopnosti hledat v datech vzory a úspěšně zachytit prostorové a časové závislosti jsou nejčastěji využívány pro zpracování, klasifikaci, rozpoznávání a segmentaci obrazů, počítačové vidění nebo rozpoznávání obličejů, své využití ale nachází i při zpracování 1D sekvencí, například při zpracování přirozeného jazyka nebo biomedicínských či audio signálů.

3.1.1 Typy vrstev používaných v CNN

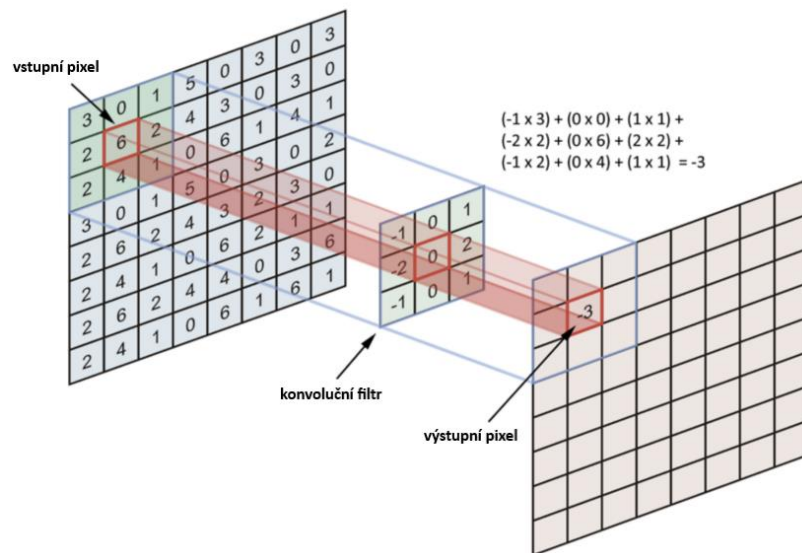
CNN je v základu tvořena třemi typy vrstev, a to jmenovitě konvolučními, poolingovými a plně propojenými vrstvami. Tyto typy vrstev mohou být doplněné o další specifické vrstvy. Architektury CNN jsou tvořeny libovolným počtem vrstev, obvykle se skládají ze sady konvolučních a poolingových vrstev naskládaných za sebe s tím, že posledních několik vrstev je plně propojených, v některých případech mohou obsahovat desítky až stovky vrstev. S rostoucím počtem vrstev se zvyšuje složitost modelu, tím roste i výpočetní náročnost.

¹⁶ doslovně může být přeloženo jako *dlouhodobá krátkodobá paměť*

¹⁷ z anglického Neural Network

Konvoluční vrstva

Prvním typem je konvoluční vrstva, která ze vstupních dat extrahuje relevantní příznaky. Každá konvoluční vrstva je složena ze sady malých filtrů, které jsou taky nazývané jako konvoluční jádra (kernely) nebo neurony. Filtr je ve své podstatě matice hodnot vah (v 1D případě vektor hodnot), která je postupně posunována po vstupních datech (ve 2D případě obraz, v 1D signál), a je počítána konvoluce mezi hodnotami filtru a částí dat, kterou filtr zrovna překrývá. Princip konvoluční vrstvy je znázorněn na obrázku 3.1. Výstup konvoluce mezi celým vstupním vzorkem a filtrem se nazývá aktivační (příznaková) mapa.



Obrázek 3.1 Princip konvoluční vrstvy (převzato a upraveno z [39])

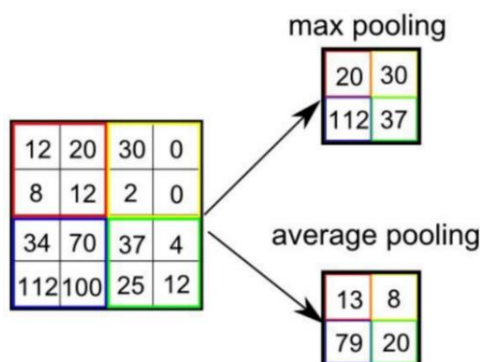
Parametry, které se dají v každé konvoluční vrstvě nastavit, jsou počet a velikost konvolučních jader, délka kroku (anglicky *stride*) a aktivační funkce (o té je řeč v kapitole 3.1.2). Počet konvolučních jader určuje počet příznakových map, které budou výstupem konvoluční vrstvy, a tím tedy i hloubku výstupu. Často se využívá většího počtu kernelů, běžně od 16 do 512 [40]. V procesu trénování sítě se hledá vhodná hodnota jejich vah tak, aby byly filtry schopny ze vstupních dat „vytáhnout“ relevantní příznaky (například lokální vzory). Rozměry konvolučních jader jsou (většinou mnohem) menší než rozměry vstupních dat, u obrazů jsou rozměry jádra typicky 3x3, 5x5 nebo 7x7, v 1D konvolučních sítích může být délka filtru i větší [41]. Délka kroku stanovuje, o kolik vzorků se konvoluční jádro posune při výpočtu konvoluce. Tím jsou ovlivněny rozměry příznakové mapy, kdy platí, že čím větší je délka kroku, tím menší bude velikost výstupní příznakové mapy.

Důležitou otázkou při využití konvoluce je, jak bude naloženo s krajními vzorky. Při normálním použití konvoluce budou rozměry výstupu menší než vstup v závislosti na

velikosti filtru. Tento problém řeší tzv. padding¹⁸ neboli přidání nul na kraje matice (vektoru). Při použití validního paddingu (anglicky *valid padding*) jsou rozměry příznakové mapy nižší než rozměry vstupních dat, při použití stejného paddingu (anglicky *same padding*) jsou rozměry vstupu i aktivačních map stejné, při použití plného paddingu (anglicky *full padding*) bude příznaková mapa větší než vstup.

Poolingová vrstva

Poolingová vrstva bývá často použita hned za konvoluční, provádí podvzorkování, snižuje tedy rozměry aktivačních map a tím i výpočetní náročnost, redukuje přeučení sítě. Poolingové vrstvy zároveň zachovávají důležité informace, které jsou pozičně invariantní vůči malým posunům, tím se zvyšuje robustnost sítě. Parametry, který se u těchto vrstev nastavují, jsou velikost filtru a délka kroku. Ovlivňují, o kolik bude zmenšena velikost aktivační mapy. Dvěma hlavními typy poolingů je max pooling a average pooling (viz obrázek 3.2). Výstupem z average pooling vrstvy je průměr všech hodnot z dané oblasti. Při max pooling je z oblasti dat, kterou pokrývá filtr, vybrána maximální hodnota. Někdy se také používá sum pooling, jehož výstupem je součet hodnot z dané oblasti příznakové mapy.

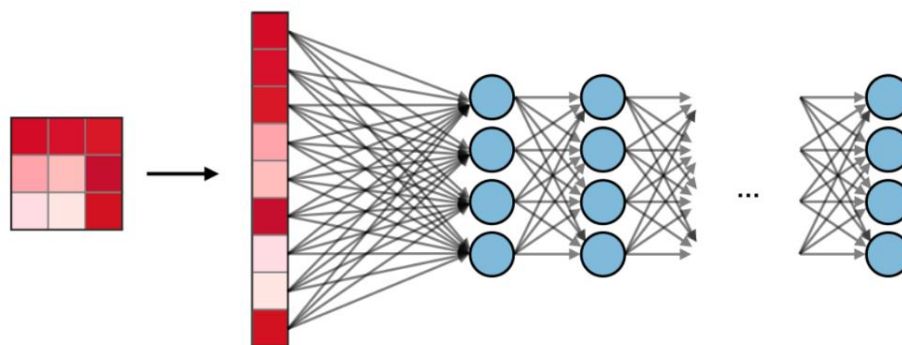


Obrázek 3.2 Příklad poolingů s filtrem velikosti 2x2 a délkou kroku 2 (převzato z [42])

Plně propojená vrstva

Plně propojené vrstvy (viz obrázek 3.3) tvoří několik posledních vrstev sítě. Termín „plně propojené“ značí, že každý neuron předchozí vrstvy je propojený s každým neuronem následující vrstvy. Slouží k naučení nelineárních kombinací všech příznaků, které jsou výstupem poslední konvoluční nebo poolingové vrstvy. Příznaky jsou před vstupem do plně propojené vrstvy obvykle zploštěny (anglicky *flattened*) neboli transformovány do 1D vektoru. Plně propojené vrstvy se typicky používají ke klasifikaci nebo predikci [39],[42],[43],[44].

¹⁸ doslovně může být přeloženo jako *vycpání*



Obrázek 3.3 Plně propojené vrstvy (převzato z [44])

3.1.2 Aktivační funkce

Aktivační funkce je matematická rovnice, která rozhoduje o výstupu jednotlivých vrstev, a tím i o výstupu celé neuronové sítě. Aktivační funkce každého neuronu rozhoduje na základě relevance váhované sumy jeho vstupních hodnot k celkové predikci modelu o tom, zda bude neuron být aktivovaný nebo ne. Pomáhá tak síti vybrat pouze důležité informace a potlačit irelevantní data. NN bez aktivační funkce je v podstatě pouze model lineární regrese.

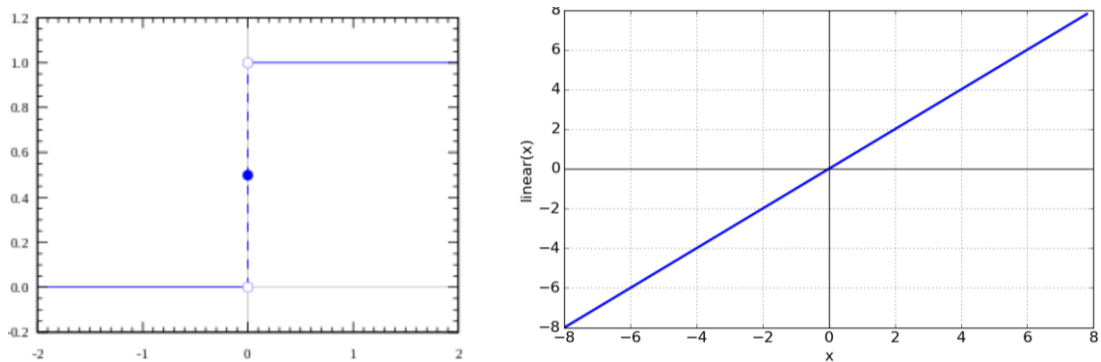
Aktivační funkce se dají rozdělit na lineární a nelineární. Moderní modely neuronových sítí používají nelineární aktivační funkce, které do sítě zavádí nelinearitu. Tyto aktivační funkce umožňují modelu vytvářet mapování mezi vstupy a výstupy, které je nezbytné pro učení a modelování komplexních dat jako jsou obrázky, video či nelineární vícedimenzionální sady dat. Nelineární aktivační funkce umožňují tvorbu hlubokých neuronových sítí, zvyšují schopnost modelu naučit se řešit komplexní problémy a dělat přesnější predikce.

Základní aktivační funkce

Mezi nejjednodušší aktivační funkce patří skoková funkce (viz obr. 3.4 vlevo). Neuron je aktivovaný v případě, kdy je hodnota vstupu x vyšší než práh p (nejčastěji 0), v opačném případě aktivován není

$$f(x) = \begin{cases} 1, & x \geq p \\ 0, & x < p \end{cases} \quad (3.1)$$

kde $f(x)$ je výstup aktivační funkce. Skoková funkce může být použita jako aktivační funkce u jednoduchého binárního klasifikátoru. Gradient skokové funkce je nulový, při využití algoritmu zpětného šíření chyby u trénování sítě se tak váhy vrstvy nemění.



Obrázek 3.4 Vlevo příklad skokové funkce (převzato z [45]), vpravo lineární funkce (převzato z [46])

Další jednoduchou aktivační funkcí je lineární funkce (viz obr. 3.4 vpravo). Její výstup $f(x)$ je proporční vstupu x

$$f(x) = a \cdot x, \quad (3.2)$$

kde a může být jakákoli konstanta. Rozsah výstupních hodnot je $(-\infty, \infty)$. Pokud jsou všechny vrstvy neuronové sítě aktivovány lineární funkcí, je celá síť ekvivalentní jedné vrstvě s lineární aktivací, protože lineární kombinace lineárních funkcí je pořád lineární funkce. NN s lineární aktivační funkcí je jednoduchý lineární regresní model, její schopnost řešit komplexní problémy je omezená. Gradient lineární funkce je konstantní (je roven hodnotě a). Změna hodnot vah při učení tak bude pořád stejná a nezávislá na hodnotě vstupu x . Pro komplexnější problémy jsou vhodnější funkce, které do sítě zavádí nelinearitu.

Sigmoidální (logistická) funkce

Sigmoidální funkce je jednou z nejpoužívanějších nelineárních funkcí, která normalizuje výstup každého neuronu, jelikož hodnoty transformuje do rozsahu 0 až 1

$$f(x) = \frac{1}{1 + e^{-x}}, \quad (3.3)$$

kde e je Eulerovo číslo. Pro velké negativní hodnoty x se výstup blíží 0, pro velké pozitivní hodnoty se naopak přibližuje k 1. Sigmoidální funkce se využívá nejčastěji při binární klasifikaci, kdy je výstupem modelu predikce pravděpodobnosti.

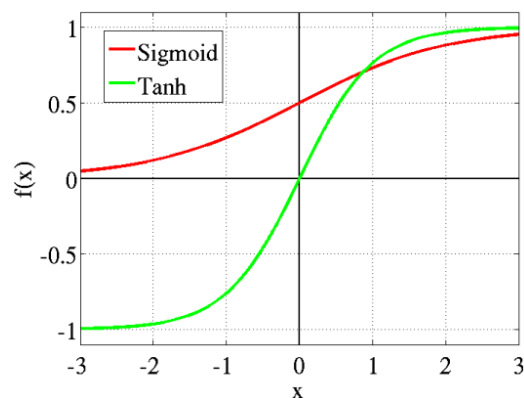
Funkce je spojitá a diferencovatelná na celém rozsahu. Gradient je však pro hodnoty x mimo rozsah od -3 do 3 velmi malý, při zvyšující se absolutní hodnotě x je prakticky nulový. Tím je do sítě zaveden problém mizejícího gradientu, kdy se kvůli extrémně malé hodnotě při trénování sítě prakticky nemění váhy, síť se tedy neučí nebo je průběh učení příliš pomalý. Sigmoidální funkce navíc není symetrická kolem počátku (viz obrázek 3.5), takže bude mít výstup všech neuronů stejné znaménko. Tento problém řeší hyperbolický tangens.

Hyperbolický tangens (*tanh*)

Funkce \tanh je velmi podobná sigmoidální funkci, ve skutečnosti je to její škálovaná verze

$$f(x) = \frac{2}{1 + e^{-2x}} - 1 = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (3.4)$$

Hyperbolický tangens je na rozdíl od sigmoidální funkce symetrický kolem počátku, rozsah hodnot je od -1 do 1 (porovnání viz obr. 3.5). Všechny ostatní vlastnosti jsou prakticky stejné jako u sigmoidální funkce, \tanh také může zavádět problém mizejícího gradientu. Další nevýhodou obou funkcí je velká výpočetní náročnost kvůli exponenciální funkci, tento problém řeší aktivační funkce ReLU.



Obrázek 3.5 Sigmoidální a \tanh funkce (převzato z [46])

ReLU

Funkce ReLU je jednou z nejvyužívanějších nelineárních aktivačních funkcí. Funkce je po částech lineární, pro pozitivní hodnoty x se vstupní hodnota rovná výstupní a výstup všech záporných hodnot je 0, funkce tak transformuje hodnoty do rozsahu $(0, \infty)$

$$f(x) = \max\{0, x\}. \quad (3.5)$$

Její výhodou oproti sigmoidální a \tanh funkci je menší výpočetní náročnost, protože pro výpočet není nutná exponenciální funkce, navíc se pro negativní vstupní hodnoty výstup rovná 0 a některé neurony tak nebudou aktivovány, což umožňuje síti rychleji konvergovat.

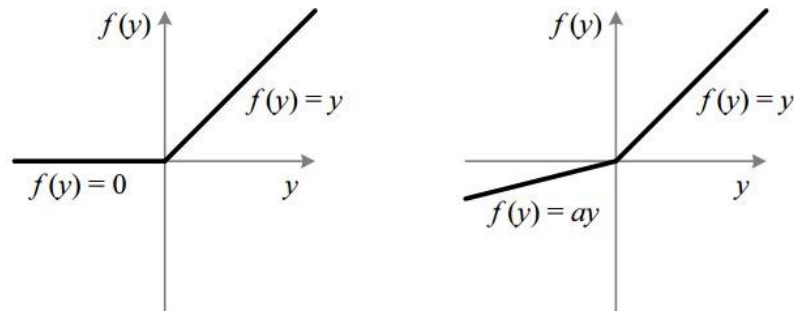
Gradient je pro záporné hodnoty x nulový, což může vést k při trénování sítě k tzv. problému umírajícího ReLU (anglicky *dying ReLU problem*). Neurony se mohou dostat kvůli nulovému gradientu do stavu, kdy se přestanou aktualizovat jejich váhy, a určitá část sítě se pak stane pasivní. Tento problém řeší různé modifikace ReLU.

Leaky ReLU

Leaky ReLU je jednou z variant, která řeší problém „mrtvých“ neuronů. Místo toho, aby byla aktivace pro záporné hodnoty nulová, je definována jako malá lineární složka x

$$f(x) = \begin{cases} x, & x \geq 0 \\ a \cdot x, & x < 0 \end{cases} \quad (3.6)$$

kde a je malá nenulová hodnota, často je rovna 0,01. Gradient záporných hodnot je tak velmi malá nenulová hodnota. Rozsah hodnot Leaky ReLU je $(-\infty, \infty)$. Výsledky při jejím použití však nejsou konzistentní. Existují i další varianty jako parametrizované ReLU či ELU¹⁹, u které je záporná část exponenciální. Rozdíl mezi funkcí ReLU a Leaky ReLU je vidět na obrázku 3.6.



Obrázek 3.6 Vlevo ukázka ReLU, vpravo Leaky ReLU (převzato z [46])

Softmax

Funkce softmax je zobecněnou verzí logistické aktivační funkce, často je popisována jako kombinace více sigmoidálních funkcí. Tato matematická funkce je využívána v poslední vrstvě sítě. Slouží ke klasifikaci do více tříd, jejím výstupem jsou normalizované pravděpodobnosti příslušnosti vzorku do každé ze tříd. U sítě, která řeší klasifikaci do více tříd, by měla poslední vrstva mít tolik neuronů, kolik je klasifikačních tříd. Součet všech pravděpodobností je roven 1 [45],[46],[47],[48].

3.1.3 Kriteriaální funkce

Kriteriaální (chybová) funkce měří predikční a generalizační schopnost modelu tak, že porovná hodnotu skutečného výsledku a výsledku predikovaného modelem. Vysoká hodnota chybové funkce značí, že se od sebe predikovaný a skutečný výsledek výrazně liší, cílem učení modelu je proto minimalizace výstupu kriteriaální funkce. Kriteriaální funkce se dají rozdělit na 2 typy podle toho, jestli se využívají k řešení regresních nebo klasifikačních úloh.

Základní kriteriaální funkce pro regresní úlohy jsou MSE²⁰ neboli střední kvadratická chyba a MAE²¹ neboli průměrná absolutní chyba. MSE je výpočet průměru kvadrátu rozdílů mezi skutečnou a predikovanou hodnotou, MAE výpočet průměrné absolutní hodnoty těchto rozdílů

¹⁹ z anglického Exponential Linear Unit

²⁰ z anglického Mean Squared Error

²¹ z anglického Mean Absolute Error

$$L_{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - p_i)^2, \quad (3.7)$$

$$L_{MAE} = \frac{\sum_{i=1}^n |y_i - p_i|}{n}, \quad (3.8)$$

kde L_{MSE} a L_{MAE} je velikost chyby, n počet vzorků, y_i je skutečná hodnota *itého* vzorku a p_i výsledek predikovaný modelem. Vzhledem k umocnění rozdílu patří mezi nevýhody MSE vysoká citlivost k odlehlým hodnotám. MAE je k šumu více robustní, její nevýhodou je však problematický výpočet derivace a více možných řešení.

Často využívanou kritériální funkcí u klasifikačních problémů je křížová entropie (CE²²), která kvantifikuje rozdíl mezi dvěma rozděleními pravděpodobnosti. V případě binární klasifikace, tedy rozdělení do dvou tříd, se používá binární křížová entropie (BCE²³)

$$L_{BCE} = -\frac{1}{n} \sum_{i=1}^n [y_i \cdot \ln(p_i) + (1 - y_i) \cdot \ln(1 - p_i)], \quad (3.9)$$

kde L_{BCE} je velikost chyby a ostatní symboly mají stejný význam jako v rovnicích (3.7) a (3.8). U klasifikace do více tříd lze využít kategoričká křížová entropie (CCE²⁴) nebo SCCE²⁵, které se liší pouze v definici skutečných hodnot y . Mezi další kritériální funkce využívané pro klasifikaci patří například Kullbackova-Leiblerova divergence [49],[50],[51],[52].

3.1.4 Optimalizační metody

Optimalizace je ve strojovém učení proces hledání parametrů sítě (vah a prahů), které minimalizují kritériální funkci. Optimalizační algoritmy (metody) se proto využívají ke snížení kritériální funkce tím, že upravují parametry sítě v závislosti na výstupu této funkce. Základním algoritmem je metoda gradientního sestupu (anglicky *gradient descent*), od které jsou odvozeny další varianty a optimalizační algoritmy.

Základní metoda gradientního sestupu

Gradient je směr růstu funkce, závisí na první derivaci chybové funkce. Parametry sítě se mění proti směru gradientu, velikost změny závisí na kroku učení. Výhodou této metody je lehký výpočet i implementace, u konvexních funkcí vždy konverguje do globálního optima, u nekonvexních funkcí do lokálního optima. Parametry se ale aktualizují po každé epoše, tzn. po výpočtu gradientu pro všechny vzorky. Metoda je tedy velmi pomalá, při

²² z anglického Cross Entropy

²³ z anglického Binary Cross Entropy

²⁴ z anglického Categorical Cross Entropy

²⁵ z anglického Sparse Categorical Cross Entropy

špatně zvoleném kroku učení je ke konvergenci nutný velký počet epoch. Je nepoužitelná pro příliš velký trénovací dataset, má velké nároky na paměť, může uváznout v lokálním minimu nebo sedlovém bodě.

Mezi varianty gradientního sestupu patří stochastický gradientní sestup (SGD²⁶) a mini-batch gradientní sestup. U SGD jsou parametry modelu upraveny po každé iteraci, tzn. po každém trénovacím vzorku. Metoda je proto výpočetně náročnější, ale konverguje rychleji než gradientní sestup a nemá tak velké nároky na paměť. Kvůli častým aktualizacím parametrů dochází k oscilacím hodnot, parametry mají velký rozptyl. Přitom není využito přímo gradientu, pouze jeho odhadu. Mini-batch varianta je z těchto tří metod nejpoužívanější, parametry modelu upravuje každých n iterací, kde n je konečný počet vzorků (velikost dávky dat). Metoda rychleji konverguje, snižuje oscilaci a rozptyl hodnot parametrů.

U těchto základních variant velmi záleží na nastavení hodnoty kroku učení, který je navíc konstantní pro všechny parametry. Velkým problémem je také uvíznutí v lokálním extrému nebo sedlovém bodě, kdy je kvůli stejné hodnotě chyby ve všech směrech nulový gradient.

Metody založené na gradientním sestupu

Problém nulového gradientu a uvíznutí v lokálním extrému částečně řeší u složitějších metod přidání hybnosti (anglicky *momentum*), která přináší informaci o předchozích gradientech. Parametr hybnost je exponenciálně váhovaný průměr gradientů z předchozích kroků. Úprava parametrů modelu tak závisí nejen na současném gradientu, ale i na gradientech předchozích. Výhodou SGD s hybností je tak rychlejší konvergence, možné řešení problému uvíznutí v lokálním extrému, snížení oscilací a rozptylu parametrů. Nevýhodou je zavedení dalšího hyperparametru, který musí být správně nastaven.

Podobnou metodou je Nesterov Accelerated Gradient (NAG), která často konverguje rychleji než SGD s hybností. Gradient je totiž počítán s ohledem na směr hybnosti v aktuální iteraci, je tak zabráněno rychlému pohybu ve špatném směru a zároveň se zvyšuje šance na nalezení lokálního či globálního minima.

Metody s adaptivním krokem učení

Adagrad (z anglického *Adaptive gradient*) používá jiný krok učení pro každý parametr. Změna kroku učení každého parametru závisí na součtu druhých mocnin všech předchozích gradientů. Výhodou je, že není nutná žádná ruční úprava kroku učení. Mezi nevýhody patří velká výpočetní náročnost, kvůli akumulaci předchozích gradientů je navíc krok učení během trénování postupně zmenšován. Při příliš malém kroku učení se pak váhy téměř nemění, a proces učení se tak může zastavit před dosažením globálního optima. Tento problém řeší metody Adadelta či RMSProp, místo akumulace kvadrátů

²⁶ z anglického Stochastic Gradient Descent

všech předchozích gradientů se jich bere v potaz pouze určitý počet.

Adam (z anglického *Adaptive Moment Estimation*), v dnešní době jeden z nejpoužívanějších algoritmů, se dá popsat jako kombinace RMSProp a SGD s hybností, zavádí adaptivní krok učení i hybnost. Kombinuje exponenciálně se snižující součet prvních i druhých mocnin předchozích gradientů, ke změně kroku učení pro každý parametr modelu využívá odhad momentů prvního a druhého řádu (průměr a rozptyl). Mezi výhody patří rychlá konvergence a malé nároky na paměť, nevýhodou je výpočetní náročnost a zavedení dalších parametrů. Mezi další pokročilejší metody patří např. Nadam, AdaMax nebo AMSGrad [53],[54],[55].

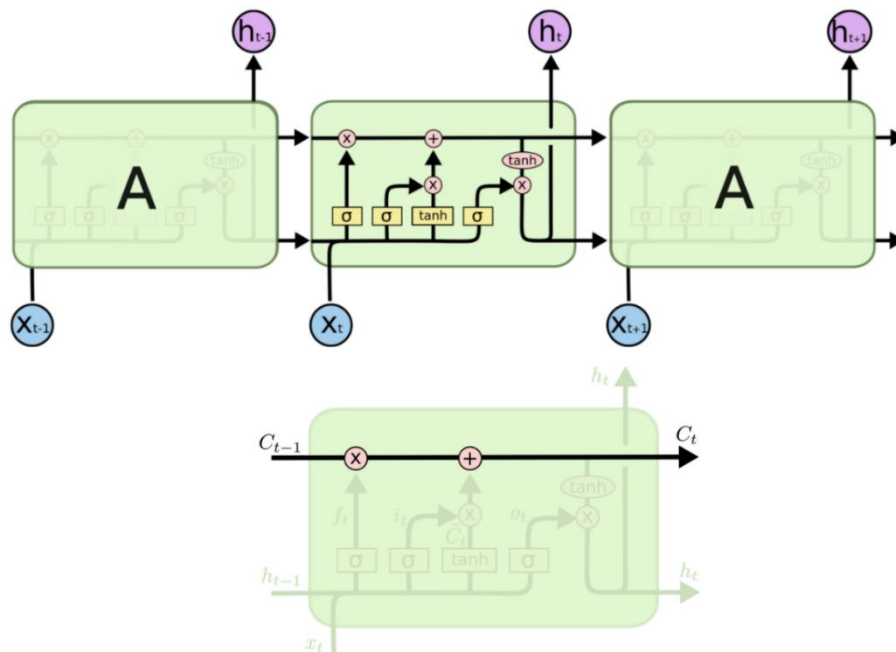
3.2 LSTM

Problémem klasických rekurentních neuronových sítí při modelování dlouhodobých závislostí je tzv. mizející gradient, kdy se v začátečních vrstvách sítě při trénování kvůli příliš malému gradientu prakticky nemění váhy, vrstvy se tak neučí. RNN si proto nedokážou zapamatovat důležité informace v delších sekvencích (nemají tzv. dlouhodobou paměť).

Tento problém řeší sítě typu LSTM, které jsou schopné naučit se dlouhodobé závislosti. Sítě LSTM jsou speciálním typem RNN sítí, byly navrženy roku 1997 Seppem Hochreiterem a Jürgenem Schmidhuberem [56], v současné době jsou užívané velmi často k různým typům úkolů jako je rozpoznávání řeči a rukopisu, generování textu nebo hudby, zpracování a predikce časových sekvencí či překlad cizího jazyka.

Struktura LSTM se dá popsat jako řetězec opakujících se modulů neuronové sítě, ve kterém modul obsahuje 4 vzájemně interagující vrstvy neuronových sítí (viz obrázek 3.7 nahoře, kde žlutý obdélník značí vrstvu neuronové sítě, růžový kruh operaci prvek po prvku, spojení dvou šipek je zřetězení neboli konkatenace dvou vektorů a rozdělení dvou šipek značí vytvoření kopie toho stejného vektoru).

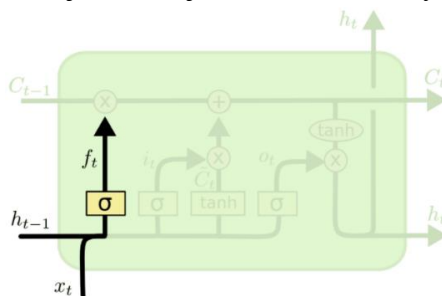
Základní součástí LSTM je stav buňky (anglicky *cell state*) C_{t-1} , viz obrázek 3.7 dole. Dá se popsat jako dlouhodobá paměť LSTM buňky. Vnitřní mechanismy, tzv. brány (anglicky *gates*), regulují tok informací LSTM buňkou, mají schopnost do stavu buňky informace přidat i je z něj odebrat. Brány se během trénování sítě učí, které informace v sekvenci jsou důležité a je nutné si je zapamatovat a přenést do dalších časových kroků, a naopak které informace nejsou relevantní a mohou být zapomenuty. Brány mají možnost propustit pouze některé informace. Jsou to jednoduché plně propojené vrstvy se sigmoidální aktivační funkcí, která hodnoty transformuje do rozsahu 0 až 1. Výstup sigmoidální funkce popisuje, které části sekvence by měly být zapamatovány (propuštěny dál). Nula značí stav, kdy se nemá dál propustit vůbec nic, naopak 1 znamená, že se má dál pustit úplně všechno. LSTM obsahuje 3 různé brány, jedná se o vstupní (anglicky *input gate*), zapomínací (anglicky *forget gate*) a výstupní bránu (anglicky *output gate*).



Obrázek 3.7 Nahoře struktura LSTM, dole stav buňky (převzato z [57])

3.2.1 Zapomínací brána

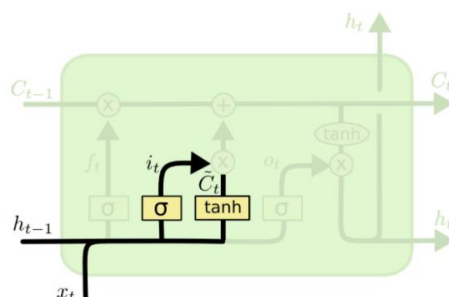
Zapomínací brána rozhoduje o tom, které předchozí informace ze stavu buňky by měly být zachovány a které mohou být zapomenuty. Jak je vidět na obrázku 3.8, do zapomínací brány vstupuje skrytý stav z předchozího časového kroku h_{t-1} (skrytý stav slouží jako krátkodobá paměť sítě, pamatuje si informace ze sekvencí, které do buňky vstupovaly v předchozích časových krocích) a současný vstupní vektor (sekvence) x_t . Výstupem je vektor hodnot f_t od 0 do 1 o stejné délce jako má stav buňky C_{t-1} .



Obrázek 3.8 Zapomínací brána (převzato z [57])

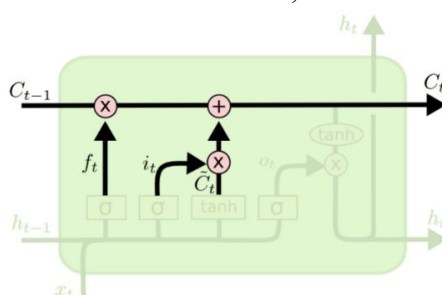
3.2.2 Vstupní brána

Vstupní brána (viz obrázek 3.9) slouží k aktualizaci dlouhodobé paměti. Nejprve je na předchozí skrytý stav a současný vstupní vektor aplikována sigmoidální funkce. Její výstup rozhodne o tom, jaké hodnoty mají být aktualizovány. Na spojený předchozí skrytý stav a současný vstup je aplikován i hyperbolický tangens, který hodnoty transformuje do rozsahu -1 a 1, což napomáhá regulaci sítě. Nakonec jsou vynásobeny výstupy z obou funkcí, kdy výstup i_t ze sigmoidy rozhoduje, které informace z výstupu C_t z tanh funkce jsou důležité a měly by tak být zachovány.



Obrázek 3.9 Vstupní brána (převzato z [57])

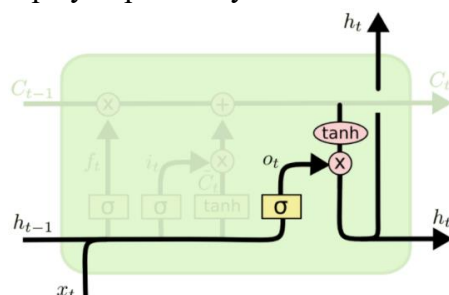
Dále je vypočítán nový stav buňky C_t . Jak je zobrazeno na obrázku 3.10, předchozí stav buňky C_{t-1} je prvek po prvku vynásoben s výstupem ze zapomínací brány f_t , a pokud jsou některé hodnoty stavu buňky násobené číslem blízkým 0, budou zapomenuty. Výstup z tohoto součinu je poté prvek po prvku sečten s výstupem vstupní brány. Stav buňky je tak aktualizován hodnotami, které LSTM buňce připadají důležité.



Obrázek 3.10 Výpočet nového stavu buňky (převzato z [57])

3.2.3 Výstupní brána

Výstupní brána rozhoduje o hodnotě nového skrytého stavu h_t , který obsahuje informace o předchozích vstupech a zároveň slouží k predikci. Nejprve je, stejně jako v obou předchozích branách, na předchozí skrytý stav a současný vstup aplikována sigmoidální funkce. Poté je na nový stav buňky vypočítaný v předchozím kroku aplikována tanh funkce, jejíž výstup je nakonec vynásoben prvek po prvku s výstupem ze sigmoidální funkce o_t . Výsledek součinu je nový skrytý stav. Nový stav buňky C_t a nový skrytý stav h_t jsou pak přeneseny do dalšího časového kroku, kde slouží jako aktuální stav buňky a aktuální skrytý stav. Princip výstupní brány lze vidět na obrázku 3.11 [57],[58],[59].



Obrázek 3.11 Výstupní brána (převzato z [57])

4. REALIZACE AUTOMATICKÉ DETEKCE KOMOROVÝCH EXTRASYSTOL

V programovém prostředí Matlab R2020a byly EKG záznamy nejprve načteny a rozděleny do trénovací, validační a testovací množiny. Následně byla v programovacím jazyce Python s využitím programu Spyder 4.2.0 a knihoven Keras a TensorFlow vytvořena a natrénována neuronová síť využívající konvolučních vrstev, LSTM jednotky a plně propojených vrstev, s jejíž pomocí byly signály klasifikovány do dvou tříd jako KES nebo normální rytmus a poté i do tří tříd, jmenovitě jako normální rytmus, KES nebo ostatní arytmie (SVES, BLRT a BPRT).

4.1 Použitá testovací databáze

MIT-BIH Arrhythmia Database (AD), pocházející ze serveru PhysioNet, zahrnuje 48 EKG záznamů²⁷, které byly pořízeny mezi lety 1975 až 1979 v Bostonu v Beth Israel Hospital²⁸ (BIH).

Ze sady více než 4000 24hodinových Holterovských EKG záznamů, ze kterých přibližně 60 % pocházelo od hospitalizovaných pacientů a 40 % od ambulantních pacientů, bylo náhodně vybráno 30 minut z 23 záznamů (očíslovaných od 100 do 124 s tím, že čísla 110 a 120 chybí). Zbývajících 25 záznamů (očíslovaných od 200 do 234, kdy nějaká čísla opět chybí) dlouhých 30 minut bylo vybráno ze stejné sady tak, aby v nich byly obsaženy méně časté, ale klinicky významné arytmie, které by v malém náhodném vzorku nemusely být zastoupené. Záznamy z první skupiny mají sloužit jako reprezentativní vzorek různorodých tvarů EKG křivky a artefaktů, se kterými by se detektor arytmií mohl při běžném klinickém používání setkat. Záznamy z druhé skupiny byly vybrány tak, aby obsahovaly komplexní komorové, junkční a supraventrikulární arytmie a abnormality v převodu vzruchu. Některé z těchto záznamů byly zvoleny díky předpokladu, že budou kvůli kvalitě signálu, rytmickým změnám a rozdílech ve tvaru QRS komplexů pro detektor představovat významný a obtížný problém.

Záznamy pochází od 25 mužů ve věku od 32 do 89 let a 22 žen ve věku od 23 do 89 let. Záznamy 201 a 202 pochází od stejného pacienta. Ve čtyřech záznamech (102, 104, 107 a 217) je přítomen rytmus řízený kardiostimulátorem.

Každý záznam obsahuje signály ze dvou svodů. Ve většině ze 48 záznamů pochází první signál z modifikovaného bipolárního končetinového svodu II (MLII²⁹), který byl získán umístěním elektrod na hrud'. Druhý signál je zpravidla modifikovaný hrudní svod V1, v několika případech V2 nebo V5 a v jednom záznamu V4. Signály byly vzorkovány frekvencí 360 Hz s rozlišením 11 bitů v rozsahu ± 5 mV.

²⁷ záznamy volně přístupné z: <https://physionet.org/content/mitdb/1.0.0/>

²⁸ nyní Beth Israel Deaconess Medical Center

²⁹ z anglického Modified Limb lead II

Záznamy byly nejprve anotovány jednoduchým QRS detektorem, který každý detekovaný úder označil jako normální. Dva kardiologové poté nezávisle na sobě anotace zkontrolovali a opravili, ke každému srdečnímu úderu přidali typ rytmu a hodnocení kvality signálu. Anotace od obou kardiologů byly porovnány a všechny nesrovnalosti vyřešeny za pomoci dalších kardiologů. Mezi lety 1980 a 2010 bylo v anotacích nalezeno a opraveno dalších 23 chyb. V databázi je tak nyní pro všech přibližně 109 tisíc srdečních úderů k dispozici anotace obsahující pozici R vlny, typ rytmu, hodnocení kvality signálu a případné další komentáře [60],[61],[62].

4.2 Vytvoření trénovací, validační a testovací množiny

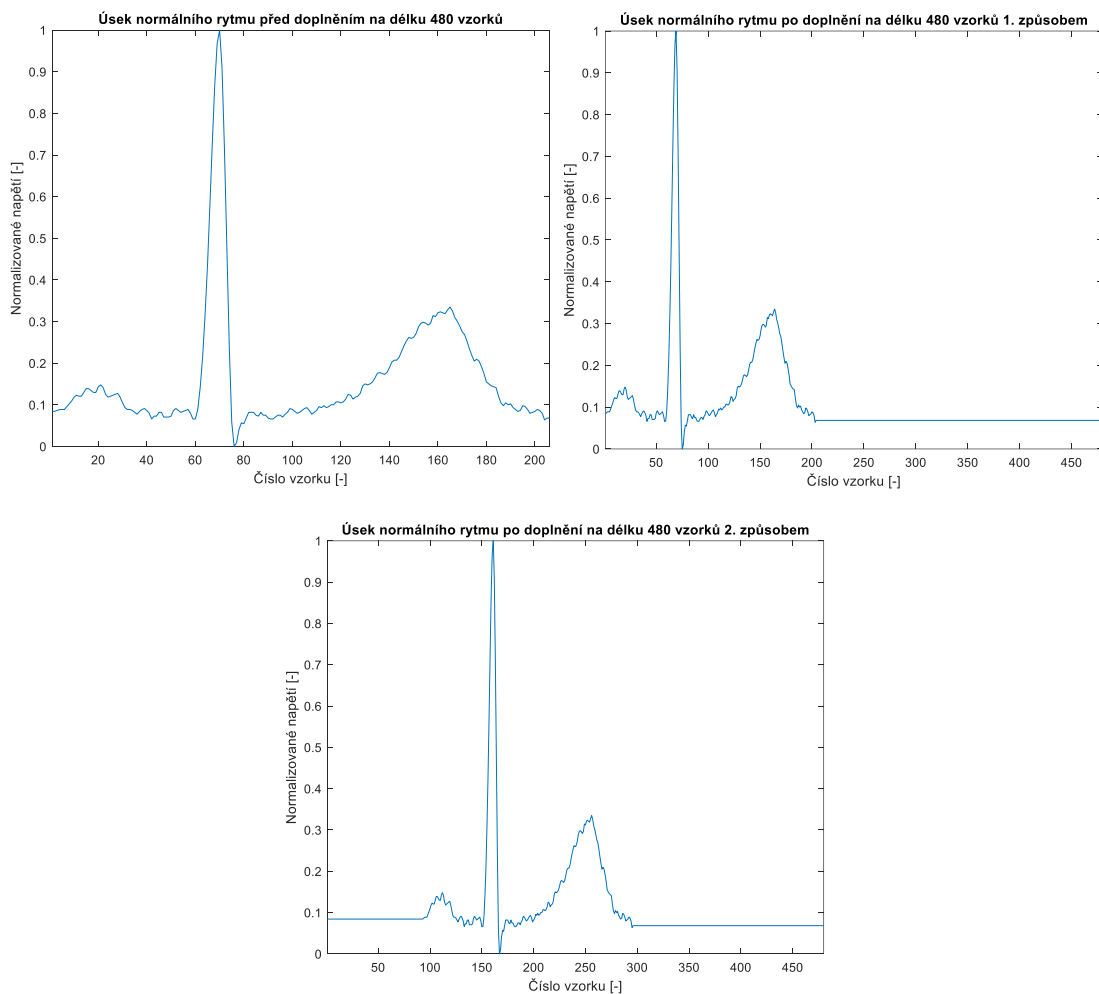
K načtení signálů a vytvoření trénovací, validační a testovací množiny byl použit program Matlab R2020a. Z MIT-BIH AD byly vybrány všechny záznamy kromě čtyř obsahujících rytmus z kardiostimulátoru (viz kapitola 4.1), celkem tak bylo použito 44 záznamů.

K načtení signálů a jejich anotací byla vytvořena funkce `nactiSignal`, ve které byly využity funkce `rdsamp` a `rdann`, které jsou součástí softwarového balíčku WFDB Toolbox pro Matlab (verze 0.10.0)³⁰. Výstupem funkce `rdsamp` jsou tři proměnné: matice složená ze dvou signálů o 650000 vzorcích, vzorkovací frekvence a vektor určující časovou osu (vzorkovací intervaly). Funkce `rdann` je použita k načtení souboru s anotacemi, zahrnuje šest výstupních proměnných. První je vektor pozic anotovaných vzorků, druhá proměnná je vektor stejné délky určující konkrétní typ anotace pro daný vzorek. Čtyři ostatní výstupní proměnné jsou další anotace, které pro následující práci nejsou podstatné. Výstupem vytvořené funkce je tedy načtený signál, pozice anotovaných vzorků a konkrétní typ anotací.

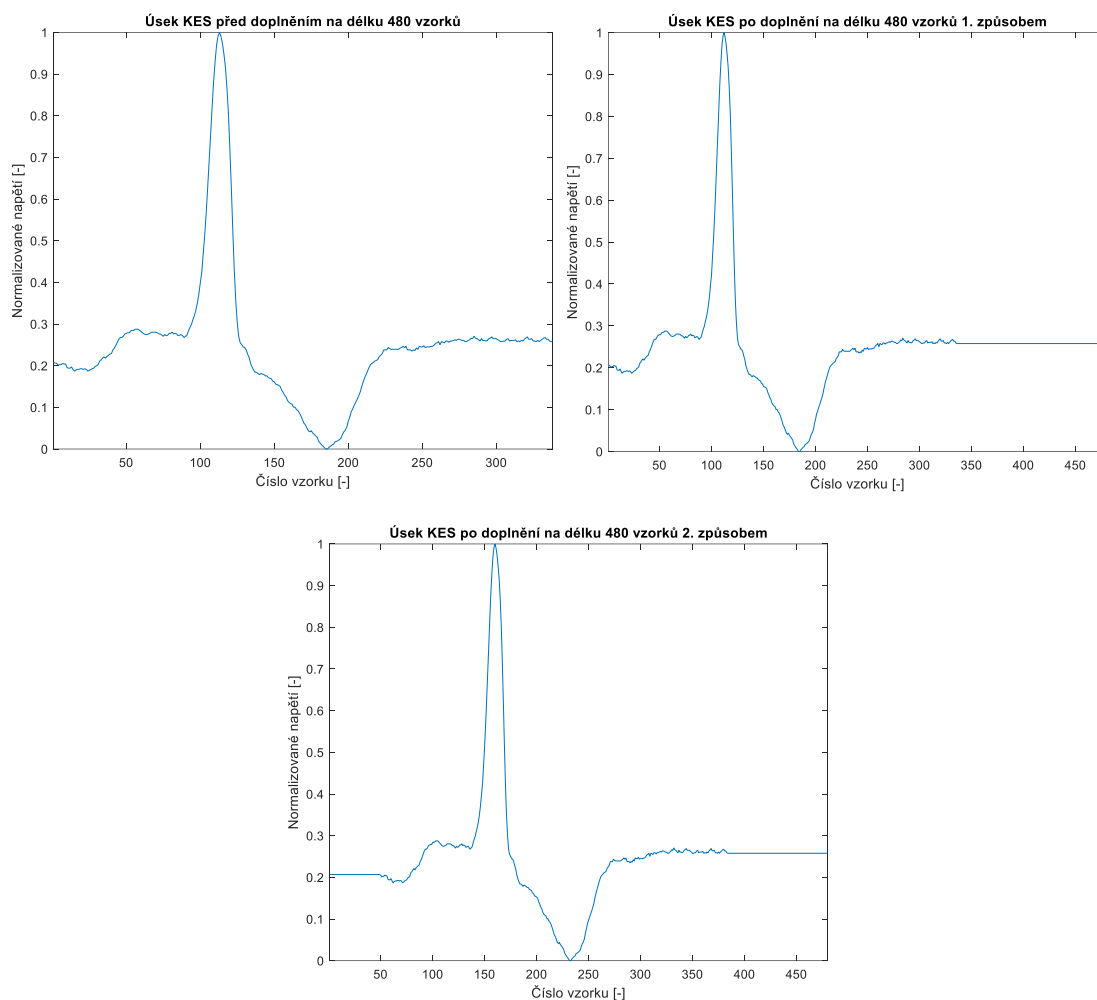
Načtené signály byly následně segmentovány na úseky odpovídající jednotlivým srdečním úderům. K tomu byla vytvořena funkce `tvorbaUseku`. Jejím vstupem jsou výstupy funkce `nactiSignal`, kýžená délka vytvořených úseků a volba způsobu doplnění úseků do této délky. V této práci byly vytvořeny úseky dlouhé 480 vzorků, což při vzorkovací frekvenci 360 Hz odpovídá úsekům o délce trvání 4/3 s, a tedy průměrné tepové frekvenci 45 tepů/min. Tato délka byla vybrána s ohledem na následující zpracování signálu pomocí neuronové sítě, a zároveň tak, aby při segmentaci jednotlivých srdečních úderů nedošlo k případné ztrátě důležité informace (např. P nebo T vln). Z každého načteného EKG záznamu byl vybrán pouze signál z prvního svodu (ve většině případů MLII, v záznamu 105 pak modifikovaný svod V5). Pomocí anotačních souborů byly vybrány pozice kmitů R u srdečních úderů s normálním rytmem (N) nebo KES, pro každý takový úder byla také spočítána délka aktuálního RR intervalu, které přibližně odpovídá době trvání daného úderu srdce. Na základě této délky pak byly segmentovány jednotlivé úseky tak, že byl vybrán počet vzorků odpovídající 1/3 délky RR intervalu před

³⁰ toolbox volně přístupný z: <https://physionet.org/physiotools/matlab/wfdb-app-matlab/>

danou R vlnou a počet vzorků odpovídající 2/3 délky RR intervalu za ní. Tímto způsobem vznikaly úseky signálu o rozdílném počtu vzorků odpovídajícím délce konkrétních RR intervalů, a bylo je tedy potřeba doplnit na jednotnou délku 480 vzorků. Doplnění bylo realizováno dvěma způsoby, u prvního způsobu byl úsek na konci doplněn na délku 480 vzorků hodnotou posledního vzorku (na obrázcích 4.1 a 4.2 vpravo nahoře), při druhém způsobu byl úsek doplněn z levé strany první hodnotou signálu a z pravé strany poslední hodnotou signálu tak, aby byl R kmit vždy na pozici 160. vzorku (na obrázcích 4.1 a 4.2 dole). Při doplnění prvním způsobem byl tak kladen důraz hlavně na délku každého srdečního úderu, u druhého způsobu pak šlo především o tvar R kmitu, který byl stále na stejné pozici. Každý úsek byl následně normalizován do rozsahu od 0 do 1 pro urychlení výpočtů a snadnější manipulaci s daty v rámci neuronové sítě za cenu ztráty informace o původní hodnotě napětí. Jak ale bylo zjištěno při testování různých nastavení sítě (viz kapitola 4.3.2), při použití původních dat bez normalizace dosahovala klasifikace horších výsledků a učení trvalo delší dobu, využití normalizace proto bylo vhodné.



Obrázek 4.1 Ukázka doplnění úseku normálního rytmu oběma způsoby: vlevo nahoře původní úsek, vpravo nahoře úsek po doplnění prvním způsobem, dole úsek po doplnění druhým způsobem (viz text)



Obrázek 4.2 Ukázka doplnění úseku KES oběma způsoby: vlevo nahoře původní úsek, vpravo nahoře úsek po doplnění prvním způsobem, dole úsek po doplnění druhým způsobem (viz text)

Výstupem funkce je matice o velikosti $M \times 480$, kde M je počet srdečních úderů typu N a KES dohromady v daném záznamu, dále vektor anotací typu rytmu o délce M , ve kterém 1 značí normální rytmus a 0 KES. V tabulce 4.1 je zaznamenáno rozdělení záznamů do trénovací, validační a testovací množiny, kde N znamená počet srdečních úderů (tedy segmentů) s normální morfologií a V značí počet úderů s KES.

Tabulka 4.1 Rozdělení dat

množina	číslo záznamu	N	V
trénovací	100, 101, 103, 105, 108, 109, 111, 112, 114, 115, 117, 118, 121, 123, 124, 200, 201, 202, 203, 205, 207, 208, 209, 214, 220, 221, 222, 223, 230, 231, 233, 234	51944	4825
validační	106, 116, 122, 210, 213	11343	1043
testovací	113, 119, 212, 215, 219, 228	11214	1034

Záznam 232 pro klasifikaci do dvou tříd nebyl využit, protože obsahuje pouze srdeční úderý typu SVES a BPRT. Celkově je tak z databáze vybráno 81403 srdečních úderů, z toho 74501 s normálním rytmem a 6902 typu KES, 8,48 % ze všech úseků tedy tvoří úseky s KES. Trénovací množina je tvořena 56769 segmenty, z toho je 8,5 % úderů typu KES. Validační množina se skládá z 12386 srdečních úderů, z nich je 8,42 % KES. Testovací množinu tvoří 12248 segmentů, z toho 8,44 % je typu KES. Trénovací množina obsahuje 69,74 % všech dat, validační 15,22 % a testovací 15,06 %.

V tabulce 4.2 jsou počty srdečních úderů s normálním rytmem a KES v každém záznamu přiblíženy detailněji, kde # značí číslo záznamu, N je počet srdečních úderů s normálním rytmem a V je počet úderů s KES.

Tabulka 4.2 Počty srdečních úderů v jednotlivých záznamech

#	N	V	#	N	V	#	N	V	#	N	V
100	2237	1	115	1952	0	202	2060	19	219	2081	64
101	1858	0	116	2301	109	203	2527	444	220	1952	0
103	2081	0	117	1533	0	205	2570	71	221	2030	396
105	2525	41	118	0	16	207	0	105	222	2060	0
106	1506	520	119	1542	444	208	1585	992	223	2028	473
108	1737	17	121	1860	1	209	2620	1	228	1687	362
109	0	38	122	2475	0	210	2421	194	230	2253	1
111	0	1	123	1513	3	212	922	0	231	314	2
112	2536	0	124	0	47	213	2640	220	232	0	0
113	1788	0	200	1742	826	214	0	256	233	2229	830
114	1819	43	201	1624	198	215	3194	164	234	2699	3

Takto vytvořené matice (trénovací o rozměrech 56769×480 , validační o rozměrech 12386×480 a testovací o rozměrech 12248×480) a jim odpovídající vektory s anotacemi srdečního rytmu byly uloženy a následně načteny do Spyderu pro další analýzu.

4.3 Realizace klasifikace EKG záznamů do dvou tříd

Ke klasifikaci rytmu na normální nebo KES v každém srdečním úderu byla využita neuronová síť, jejíž architektura byla inspirována článkem [63], který vyšel roku 2019 v časopise *Computer Methods and Programs in Biomedicine*. V této publikaci byl nejprve natrénován autoenkodér, tedy typ neuronové sítě, která vstupní data převede do příznaků v latentním (skrytém) prostoru, a poté se snaží vstupní data ze skrytých příznaků rekonstruovat tak, aby byla mezi vstupními a výstupními daty co nejmenší chyba (aby si

data byla co nejpodobnější). Při převodu do latentní reprezentace jsou vstupní data komprimována tak, že jsou rozměry vypočtených příznaků mnohem menší než rozměry vstupních dat. Autoenkodér je složen ze dvou funkčních celků. Prvním je enkodér, který data zakóduje do příznaků, druhý je dekodér, který z příznaků rekonstruuje vstupní signál. Autoenkodér slouží k výpočtu a extrakci relevantních příznaků, které signál dobře popisují, jelikož z nich dekodér dovede původní signál věrně rekonstruovat, jeho významnou vlastností je také schopnost redukce šumu. Ke klasifikaci byl poté použit již naučený enkodér, za který byla přidána LSTM jednotka a plně propojená vrstva.

Během tvorby této práce byly testovány dva přístupy. První byl stejný jako v právě popsaném článku, tedy nejprve trénování autoenkodéru, poté trénování klasifikátoru za použití již natrénovaného enkodéru; při učení se tak upravovaly pouze váhy u LSTM jednotky a plně propojené vrstvy. U druhého přístupu bylo vynecháno trénování autoenkodéru. Klasifikátor tak měl stejnou architekturu jako u prvního přístupu, ale při učení se měnily váhy všech vrstev. Jelikož klasifikace za použití druhého přístupu dosahovala lepších výsledků, je v práci tento přístup využit a popsán.

4.3.1 Architektura sítě

V rámci práce bylo testováno několik architektur neuronových sítí lišících se rozměry a rozsahem dat vstupujících do první vrstvy, celkovým počtem a rozložením vrstev a množstvím neuronů v jednotlivých vrstvách.

Tabulka 4.3 Architektura sítě

		vrstva	VF	AF	#	VV
bottleneck	1	Conv1D	8×5	ReLU	48	480×8
	2	MaxPooling1D	2	-	0	240×8
	3	Conv1D	16×5	ReLU	656	240×16
	4	BatchNorm	-	-	64	240×16
	5	MaxPooling1D	2	-	0	120×16
	6	Conv1D	64×5	ReLU	5184	120×64
	7	BatchNorm	-	-	256	120×64
	8	MaxPooling1D	2	-	0	60×64
	9	Conv1D	32×3	ReLU	6176	60×32
	10	Conv1D	1×3	ReLU	97	60×1
	11	MaxPooling1D	2	-	0	30×1
klasifikační část	12	LSTM	32	-	4352	30×32
	13	Flatten	-	-	0	960
	14	Dense	64	ReLU	61504	64
	15	Dropout	-	-	0	64
	16	Dense	1	-	65	1

S ohledem na výsledky dosažené u trénovací a validační množiny byla ke klasifikaci typu rytmu v EKG záznamech nakonec použita neuronová síť o 16 vrstvách, z toho prvních 11 vrstev tvoří tzv. bottleneck, který z původních úseků o délce 480 vzorků vytvoří vektor 30 příznaků. Příznakový vektor dále vstupuje do LSTM vrstvy a dvou plně propojených vrstev. Architektura je znázorněna v tabulce 4.3, VF znamená velikost filtru (počet jednotek), AF je aktivační funkce, # značí počet trénovaných parametrů (vah) a VV je velikost výstupu.

Vstupem do sítě jsou jednotlivé úseky srdečního cyklu o rozměrech 480×1 . Funkce některých vrstev je podrobněji vysvětlena v kapitole 3.1.1. Vrstva *Conv1D* provádí 1D konvoluci. První číslo ve sloupci VF značí počet filtrů (konvolučních jader), určuje tedy druhou dimenzi výstupního signálu (zaznačeno ve sloupci VV), první dimenze je určena délkou signálu. Druhé číslo velikosti filtru je velikost konvolučního jádra neboli délka okna, ve kterém je počítána konvoluce. Padding je u všech konvolučních vrstev nastaven na *same*, délka signálu se tedy při průchodu těmito vrstvami nemění. U všech konvolučních vrstev je využita aktivační funkce ReLU. Vrstva *MaxPooling1D* slouží k podvzorkování. V síti je u všech těchto vrstev nastavena velikost filtru i délka kroku na hodnotu 2, každá poolingová vrstva tak signál zmenšila na polovinu jeho předchozí délky. Vrstva *BatchNorm* slouží k normalizaci, pomáhá ke snížení rizika přeučení sítě. Vstupní data transformuje tak, že jejich průměr se rovná 0 a směrodatná odchylka 1. Jak již bylo řečeno, prvních 11 vrstev tvoří tzv. bottleneck, kde se se snižujícím rozměrem příznakových map zvyšuje jejich počet (hloubka). Výstupem 11. vrstvy je vektor 30 příznaků.

V síti je použito 32 LSTM jednotek (detailněji popsané v kapitole 3.2). Parametr *return_sequences* je zde nastaven na *True*, což znamená, že výstupem z této vrstvy jsou skryté stavy z každého časového kroku. Jinak by byl výstupem pouze poslední skrytý stav. Vrstva *Flatten* slouží ke změně tvaru dat, kdy je vícedimenzionální signál transformován (zploštěn) do 1D vektoru. Vrstva *Dense* je plně propojená vrstva, počet filtrů určuje velikost výstupního signálu. Aktivační funkce první plně propojené vrstvy je ReLU. Vrstva *Dropout* v každém kroku během trénování „vypne“ určité procento náhodně vybraných neuronů. To znamená, že bude některým neuronům přidělena nulová váha, a jejich výstup tedy bude taky nulový. Tato vrstva slouží stejně jako *BatchNorm* ke snížení rizika přeučení sítě. Počet neuronů, které budou vypnuty, je stanovený parametrem *rate*. Optimální nastavení hodnoty tohoto parametru bylo testováno (viz kapitola 4.3.2), jako nejvhodnější byla nakonec vybrána hodnota 0,2. Dropout vrstva tedy ovlivnila 20 % náhodně vybraných neuronů. Poslední vrstva je plně propojená, obsahuje pouze jeden neuron. Jako aktivace byla zvolena sigmoidální funkce, jelikož vrstva slouží k binární klasifikaci (klasifikaci do dvou různých tříd – normální rytmus nebo KES).

4.3.2 Výběr vhodných hyperparametrů

Po zvolení optimální architektury byla celá síť trénována s různým nastavením čtyř

hyperparametrů, a to kroku učení, počtu epoch, velikosti dávky dat (*batchsize*) a parametru *rate* v dropout vrstvě, dále s různou normalizací vstupních dat. Pro tyto hyperparametry byly nejprve odhadem určeny rozsahy možných hodnot s ohledem na informace vyčtené z odborné literatury týkající se klasifikačních neuronových sítí, poté byly zkoušeny různé kombinace těchto hodnot. Hyperparametry, které zůstávaly i přes různé nastavení výše zmíněných parametrů stejné, byla optimalizační funkce, chybová funkce a metrika posuzující výkon modelu. Jako optimalizační funkce byl zvolen algoritmus Adam, jako chybová funkce byla vybrána binární křížová entropie a jako metrika byla zvolena přesnost.

Při trénování sítě byl použit tzv. callback, což je sada funkcí, které se během trénování dají použít, patří mezi ně například průběžné ukládání modelu, zastavení trénování, pokud se model dále neučí či průběžná úprava kroku učení. Použitý callback je vidět na obrázku 4.3, do souboru *checkpointVahy.h5* tak byly po každé epoše uloženy aktuální váhy modelu, když bylo v dané epoše na validační množině dosaženo dosud nejvyšší přesnosti.

```
# nastavení callback
modelCallback = keras.callbacks.ModelCheckpoint(
    'checkpointVahy.h5',
    save_weights_only=True,
    monitor='val_accuracy',
    mode='max',
    save_best_only=True)
```

Obrázek 4.3 Část kódu věnující se nastavení callback funkce

Jak už bylo zmíněno v kapitole 4.2, vstupní data jsou nevyvážená, jelikož je segmentů s normálním rytmem mnohem více než segmentů s KES. Při trénování sítě byl proto parametr *class_weight* nastaven tak, aby přesně odpovídal poměru KES a normálních segmentů v trénovacích datech. Znamená to, že pokud síť chybně klasifikuje segment s KES, chyba bude daleko větší než při chybné klasifikaci normálního rytmu. Tento parametr byl vypočítán za využití funkce *compute_class_weight*, která lze najít v modulu Utilities for Developers knihovny Scikit-learn.

Trénovací i validační data byla před vstupem do sítě převedena na tenzor a vždy před každou epochou náhodně promíchána, aby nedošlo k nežádoucímu učení souvislostí mezi po sobě následujícími vzorky. Jednalo se o typ učení s učitelem, jako požadovaný výstup sloužily anotace dostupné ke každému záznamu. Na obrázku 4.4 je vidět část kódu, která spouští trénování modelu.

```
vysledekUceni = model.fit(trenovaciTensor, trenovaciLabel, epochs = pocetEpoch,
    batch_size = velikostBatch, verbose = 1,
    validation_data = (validacniTensor, validacniLabel),
    class_weight = vaha, callbacks = [modelCallback])
```

Obrázek 4.4 Část kódu věnující se nastavení trénování sítě

Úspěšnost trénování byla porovnáována podle výsledků na validační množině dat, kdy bylo vždy spočítáno množství správně a chybně klasifikovaných segmentů, a z nich celková senzitivita, specificita a přesnost (viz kapitola 5.1).

Normalizace vstupních dat

S ohledem na normalizaci vstupních dat byly zkoušeny 3 různé přístupy. U prvního úseky nebyly normalizovány vůbec, aby nedošlo ke ztrátě informace o původním napětí EKG záznamů. Při tomto přístupu se výrazně zvýšil výpočetní čas, přesnost klasifikace validačních dat byla oproti trénovacím nižší kvůli velkému množství úseků KES, které byly sítí označeny jako normální. Výsledky na validační množině byly zároveň horší než u normalizovaných dat. Druhý přístup spočíval v normalizaci trénovacích, validačních a testovacích dat do rozsahu 0 až 1 tzv. min-max normalizací

$$x_{i,\text{norm}} = \frac{x_i - \min(x_{\text{tren}})}{\max(x_{\text{tren}}) - \min(x_{\text{tren}})}, \quad (4.1)$$

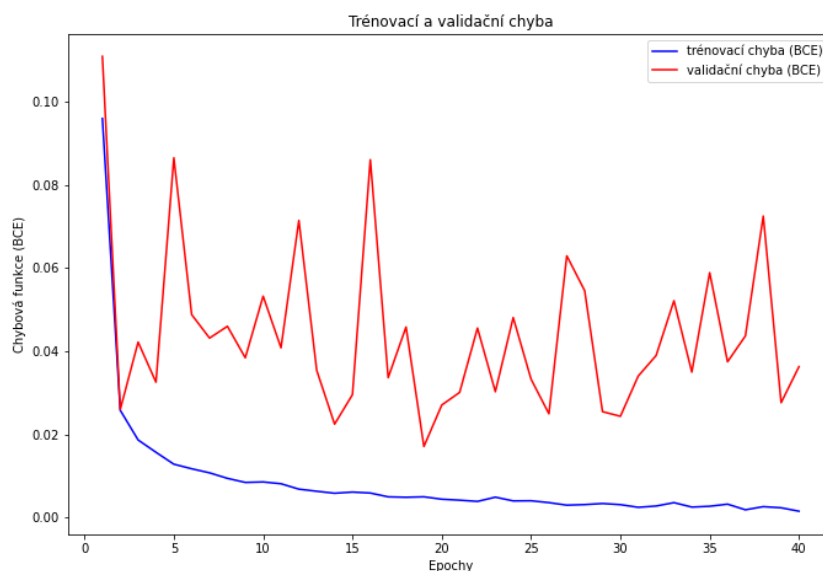
kde x_i je i tý vzorek v každém segmentu, x_{tren} značí celou trénovací množinu dat a $x_{i,\text{norm}}$ je i tý vzorek daného segmentu po normalizaci. Data tak byla normalizována pomocí minima a maxima vypočteného ze všech trénovacích dat. To ale vedlo kvůli odlehilým hodnotám k tomu, že byla většina segmentů normalizována do rozsahu přibližně 0,45 až 0,55. Jako nejvhodnější přístup byla zvolena min-max normalizace, při které bylo vypočteno maximum a minimum pro každý segment zvlášť, a všechny segmenty tak byly normalizovány do rozsahu 0 až 1 samostatně.

Krok učení

Parametr krok učení byl testován v rozsahu hodnot od 0,001 po 0,0001 s krokem 0,0001. Při nižších hodnotách kroku učení bylo ke trénování nutné využít více epoch a zvýšil se tak výpočetní čas, u validačních dat byl také vyšší počet úseků normálního rytmu, které síť klasifikovala jako KES. Při vyšších hodnotách kroku učení se síť naučila během jedné epochy, a poté se trénovací přesnost již měnila pouze nepatrně, nebo během dalších epoch naopak výrazně kolísala. Jako optimální hodnota proto byl zvolen krok učení 0,0004.

Velikost dávky dat

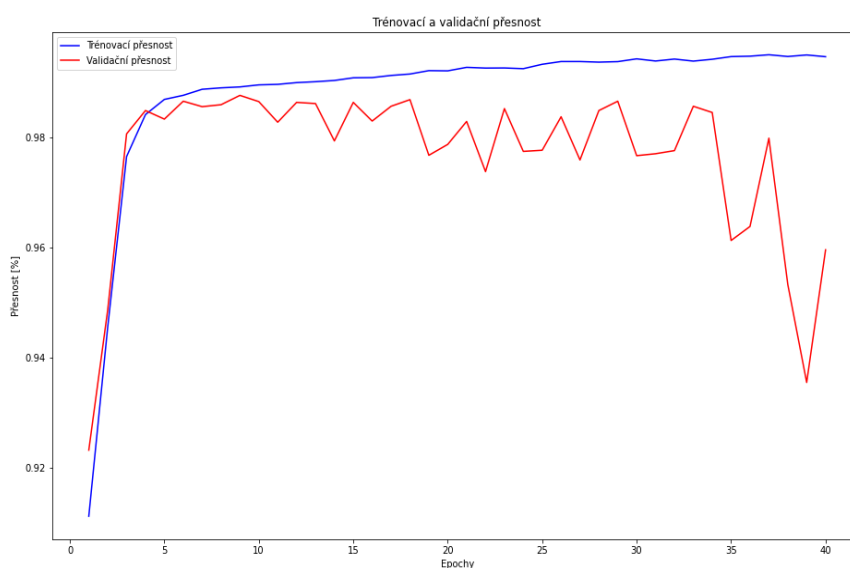
U parametru batchsize byly testovány hodnoty 16, 32, 64, 128, 256, 512 a 1024. U velkých dávek dat se u validační množiny objevovalo velké množství špatně klasifikovaných úseků vždy jednoho typu – buď normální úseky klasifikované jako KES nebo naopak. U malých dávek dat v průběhu trénování výrazně kolísala přesnost klasifikace a tím i validační chyba, jak je vidět na obrázku 4.5. Velikost dávky dat proto byla stanovena na 128, při tomto nastavení dosahovala klasifikace nejlepších výsledků.



Obrázek 4.5 Ukázka kolísání chybové funkce a přeučení sítě, velikost dávky = 32

Počet epoch

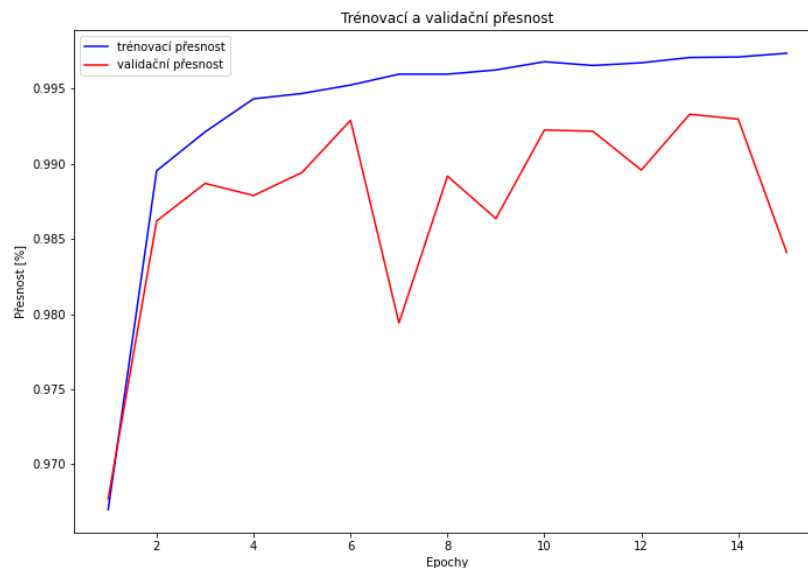
Počet epoch souvisí mimo jiné s nastavením kroku učení. Při testování různých nastavení hyperparametrů byl nejprve nastaven počet epoch na 60 a k tomu callback *Early Stopping*, který trénování zastavil předčasně, pokud se v pěti po sobě následujících epochách nezmenšila validační chyba. Trénování ale v tomto případě trvalo zbytečně dlouho a často docházelo k přeučení sítě, ukázka je na obrázku 4.6. Při nastavení kroku učení na 0,0004 a velikosti dávky dat na 128 se jako vhodný počet epoch jevila hodnota 15, jelikož při tomto počtu nedojde k přeučení sítě, zároveň se síť natrénuje dostatečně dobře.



Obrázek 4.6 Ukázka přeučení sítě při vyšším počtu epoch

Parametr rate v dropout vrstvě

Parametr *rate* ovlivňuje v každé trénovací epoše množství neuronů, které se účastní trénování sítě. Byly testovány hodnoty od 0,1 do 0,4 s krokem 0,05. U nižších nastavených hodnot tohoto parametru docházelo u sítě k přeučení, kdy u validační množiny docházelo k velkému množství převážně falešně pozitivních klasifikací. Příklad je vidět na obrázku 4.7, kdy se parametr *rate* rovnal 0,1, krok učení 0,0003 a velikost dávky 64. Při příliš velkých hodnotách parametru *rate* pak klesala přesnost i u trénovací množiny. Tento parametr byl proto nastaven na hodnotu 0,2, protože při ní bylo dosahováno nejlepších výsledků pro trénovací i validační množinu.



Obrázek 4.7 Ukázka trénování sítě při špatném nastavení hyperparametrů

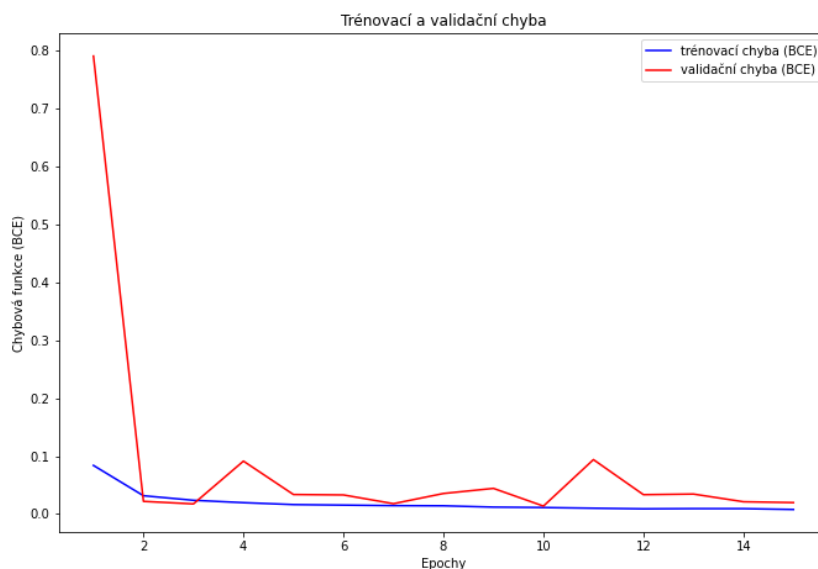
4.3.3 Trénování sítě

Po nalezení vhodných hyperparametrů, zobrazených v tabulce 4.4, byla síť trénována na datech doplněných prvním i druhým způsobem (viz kapitola 4.2) s cílem dosáhnout co nejlepších výsledků pro trénovací, validační i testovací data a také zjistit, jestli bude mezi výsledky pro oba způsoby doplnění nějaký rozdíl, a případně v čem.

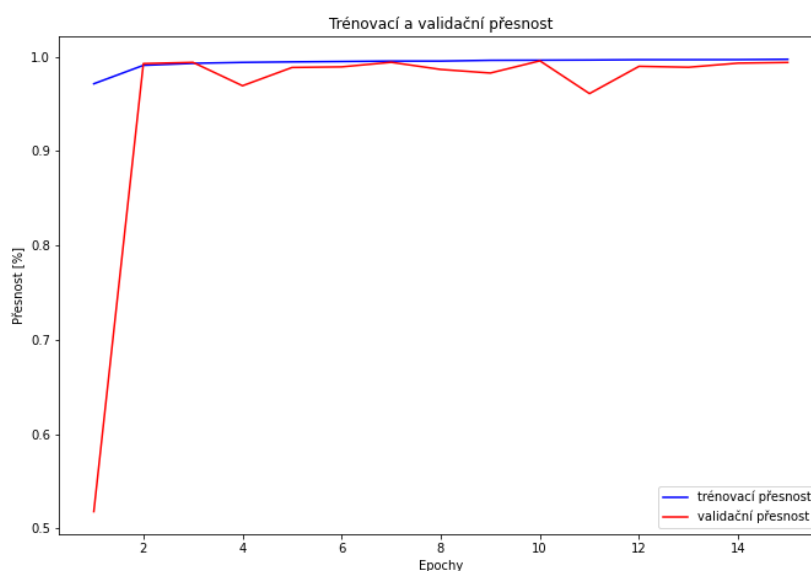
Tabulka 4.4 Nastavení hyperparametrů sítě pro klasifikaci do dvou tříd

optimalizační funkce	Adam
chybová funkce	binární křížová entropie
metrika	přesnost
krok učení	0,0004
počet epoch	15
batchsize	128
rate (dropout)	0,2

Průběh trénování pro data připravená prvním způsobem (doplnění zprava do délky 480 vzorků poslední hodnotou) je znázorněn na obrázcích 4.8 a 4.9. Jak lze pozorovat, validační chyba klesla na přibližně stejnou hodnotu jako chyba trénovací již při druhé epoše, v dalších epochách už pouze kolísala na přibližně stejné hodnotě (kromě větších výkyvů ve čtvrté a jedenácté epoše).



Obrázek 4.8 Průběh chybové funkce při trénování sítě, data doplněná prvním způsobem

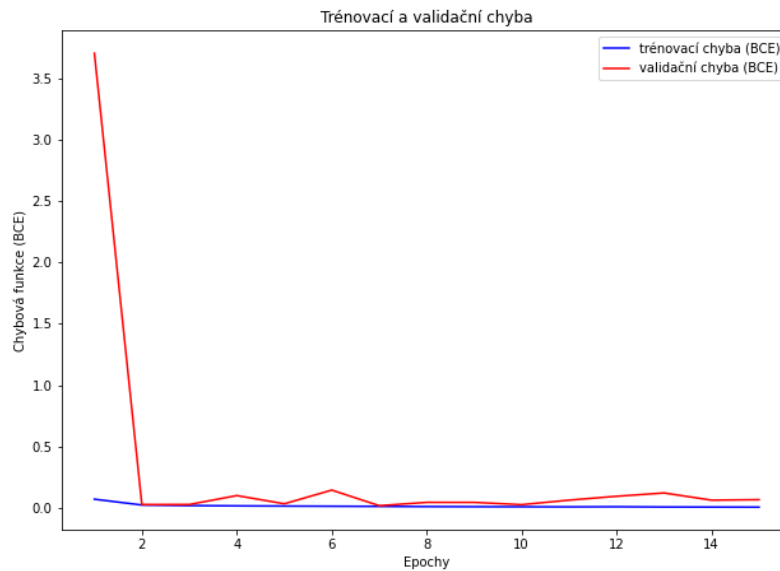


Obrázek 4.9 Průběh trénovací a validační přesnosti při trénování sítě, data doplněná prvním způsobem

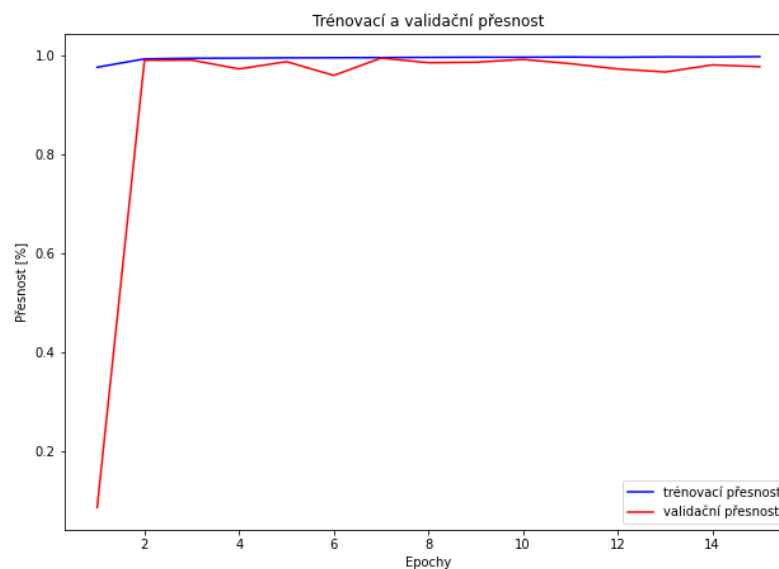
Nejvyšší přesnosti na validační množině bylo dosaženo v 10. epoše, a díky dříve zmíněnému nastavení callbacku (viz obrázek 4.3) byly uloženy váhy modelu z této epochy. Ty byly následně do modelu zpětně načteny a takto naučená síť byla uložena.

Poté byla vyhodnocena celková úspěšnost klasifikace na trénovacích, validačních i testovacích datech. Přesnost klasifikace pro trénovací i validační záznamy se rovnala 99,57 %, pro testovací pak 99,40 %. Velikost chybové funkce (binární křížové entropie) pro trénovací data vyšla 0,0121, pro validační data 0,0137 a pro testovací 0,0219. Dosažené výsledky jsou podrobněji rozebrány v kapitole 5.2.1.

Průběh trénování pro data připravená druhým způsobem (kmit R na pozici 160. vzorku, doplnění zleva i zprava první a poslední hodnotou na celkovou délku 480 vzorků) je znázorněn na obrázcích 4.10 a 4.11.



Obrázek 4.10 Průběh chybové funkce při trénování sítě, data doplněná druhým způsobem



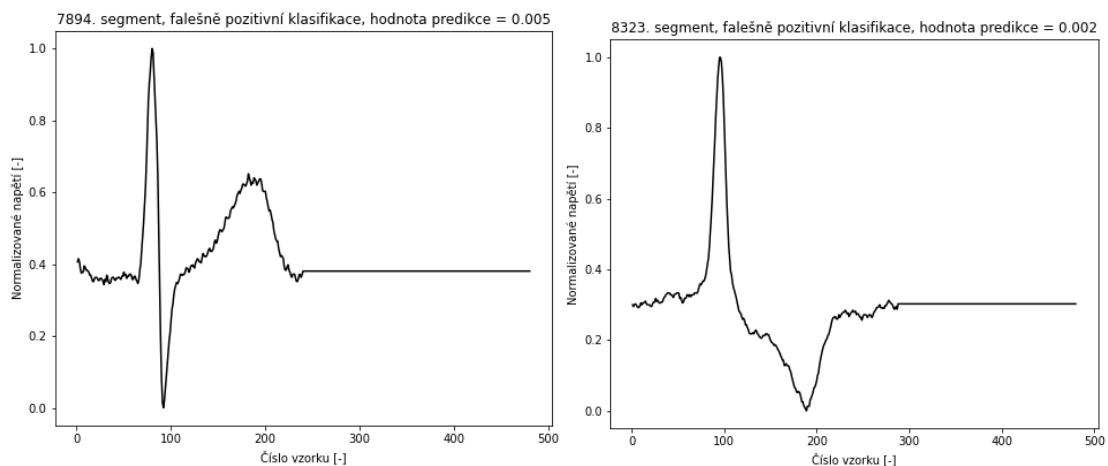
Obrázek 4.11 Průběh trénovací a validační přesnosti při trénování sítě, data doplněná druhým způsobem

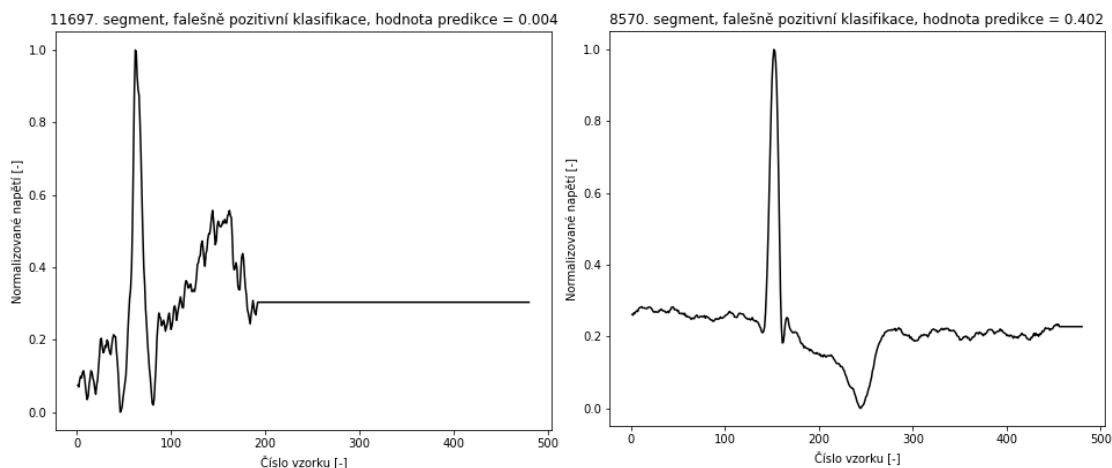
Validační chyba – stejně jako u dat připravených prvním způsobem – klesla na přibližně stejnou hodnotu jako chyba trénovací již v druhé epoše, poté už pouze kolísala na přibližně stejné hodnotě (kromě výkyvů v 6. a 13. epoše). Nejvyšší přesnosti na validační množině bylo dosaženo v 7. epoše, model s vahami z této epochy byl uložen. Přesnost klasifikace pro trénovací záznamy dosáhla 99,54 %, pro validační záznamy se rovnala 99,53 % a pro testovací 99,41 %. Velikost BCE pro trénovací data vyšla 0,0147, pro validační data 0,0178 a pro testovací 0,0221. Z těchto výsledků (a výsledků detailněji popsanych v kapitole 5.2.1) je vidět, že obě sítě naučené na různých připravených datech dosáhly přibližně stejných výsledků.

4.3.4 Ukázky výsledků klasifikace

Na následujících obrázcích jsou zobrazeny segmenty z testovací množiny dat, které byly chybně klasifikovány jako rytmus KES, když byl jejich rytmus ve skutečnosti normální (viz obrázek 4.12 a 4.14 pro data připravená prvním a druhým způsobem) a segmenty KES, které byly klasifikátorem chybně označeny jako normální rytmus (viz obrázek 4.13 a 4.15 pro data připravená prvním a druhým způsobem). Všechny signály z testovací množiny dat pochází ze svodu MLII. U segmentů byly také vypsány výstupy poslední vrstvy, tedy predikce pravděpodobnosti příslušnosti k normálnímu rytmu. Při přípravě datových množin byly segmenty normálního rytmu anotovány jako 1, výstup poslední vrstvy s hodnotou $\geq 0,5$ byl tedy klasifikován jako normální rytmus. Jako segmenty s KES byly označeny ty úseky, u kterých byl výstup poslední vrstvy menší než 0,5. Čím více se hodnota predikce přibližuje 0,5, tím méně si je klasifikátor jistý, o jaký rytmus se jedná.

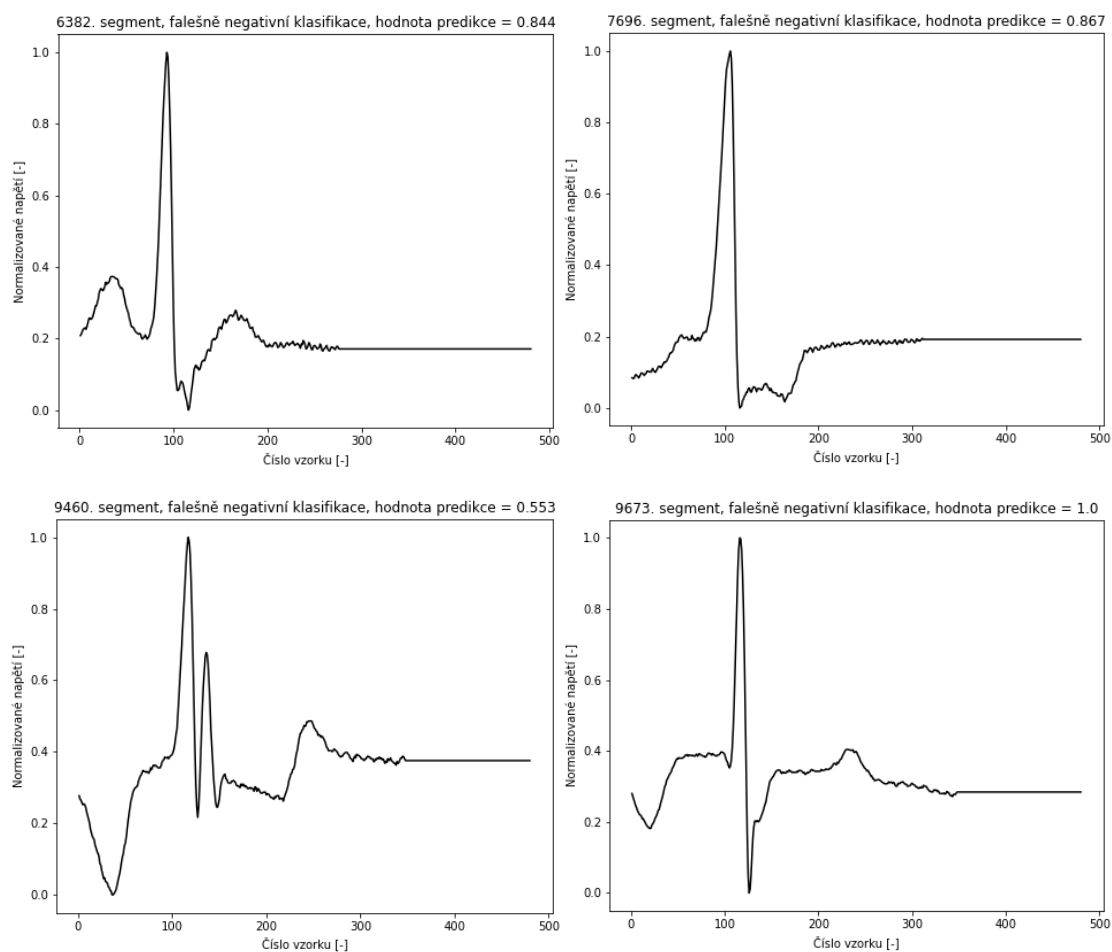
Jak je vidět u prvních tří segmentů na obrázku 4.12, hodnota predikce je $\leq 0,005$, což znamená, že byly tyto úseky klasifikátorem označeny jako KES s velkou jistotou. Hodnota predikce u posledního úseku je 0,402, klasifikátor tedy úseku přiřadil pravděpodobnost příslušnosti k rytmu KES přibližně 60 % a normálnímu rytmu 40 %.

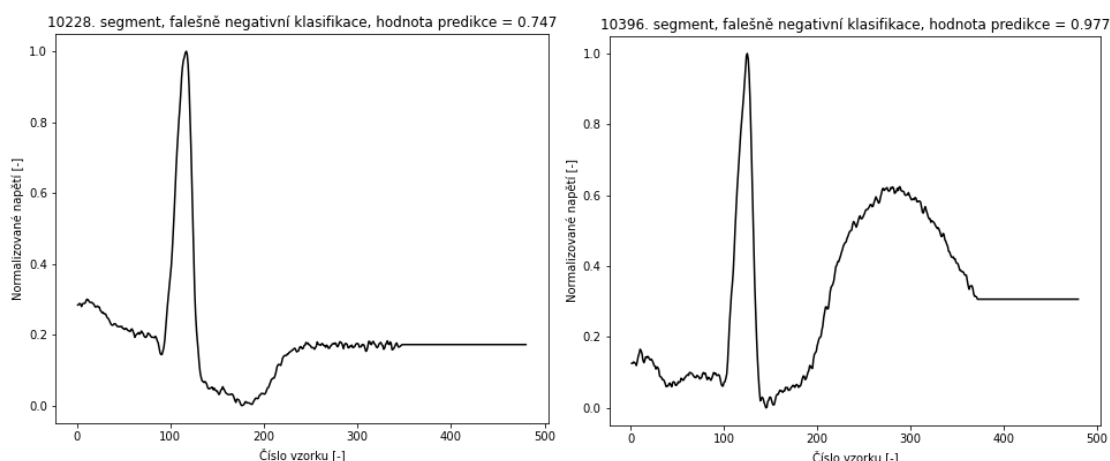




Obrázek 4.12 Ukázka falešně pozitivní klasifikace segmentů (data doplněna prvním způsobem)

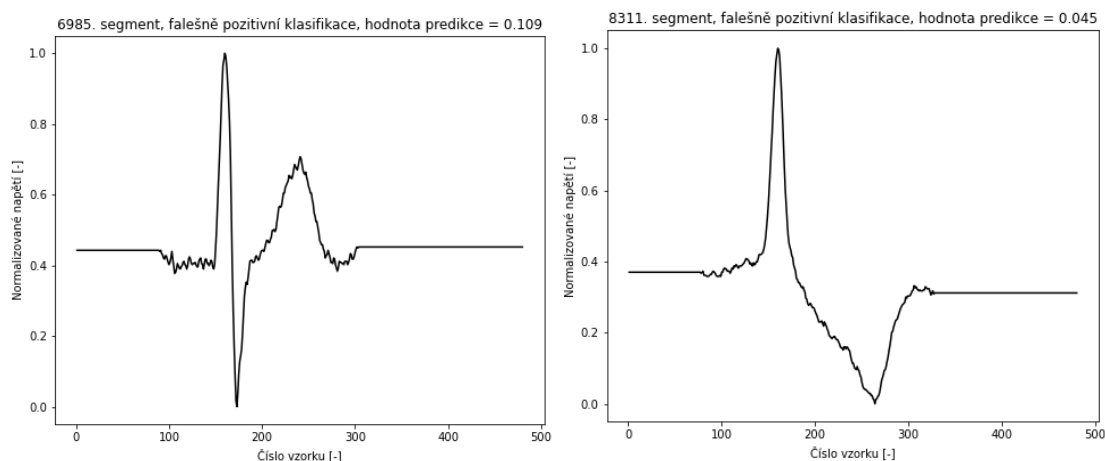
U 3. úseku na obrázku 4.13 je hodnota predikce 0,553, tedy téměř na hranici nejistoty. Naopak u čtvrtého úseku je hodnota predikce 1,0, síť tedy se 100 % jistotou tento úsek označila jako normální rytmus, i když je to podle dostupných anotací segment KES.





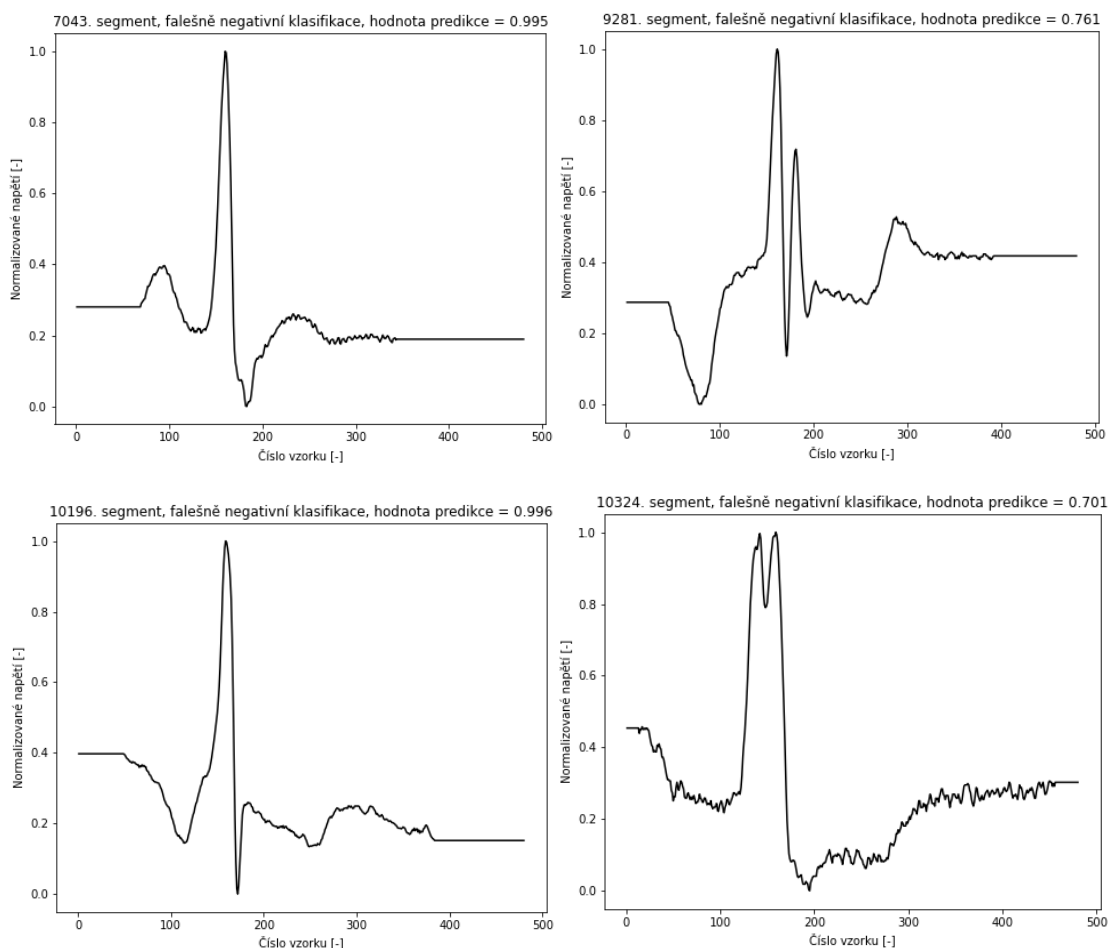
Obrázek 4.13 Ukázka falešně negativní klasifikace segmentů (data doplněna prvním způsobem)

Jak je vidět při porovnání obrázků 4.12 a 4.14, obě sítě špatně klasifikovaly stejný typ srdečních úderů anotovaných databází jako normální rytmus. Vzhledem k průběhu signálu ve druhém segmentu ale vyvstává otázka, zda jsou všechny srdeční údery v MIT-BIH AD opravdu anotovány správně.



Obrázek 4.14 Ukázka falešně pozitivní klasifikace segmentů (data doplněna druhým způsobem)

Stejně jako u výše zmíněné dvojice obrázků se dají porovnat i průběhy křivek segmentů na obrázcích 4.13 a 4.15. Závěr je v tomto případě obdobný, obě natrénované sítě většinou nesprávně klasifikovaly ty stejné segmenty, a způsob doplnění dat tak na výsledek neměl příliš velký vliv.



Obrázek 4.15 Ukázka falešně negativní klasifikace segmentů (data doplněna druhým způsobem)

4.4 Klasifikace EKG záznamů do tří tříd

Neuronová síť s téměř stejnou architekturou byla následně trénována ke klasifikaci srdečního rytmu v EKG záznamech do tří kategorií: normální rytmus, KES a ostatní arytmie (síňové extrasystoly a blokáda levého i pravého raménka Tawarova). Jednotlivé úseky byly připraveny stejným způsobem jako v kapitole 4.2, délka úseku tedy byla určována podle aktuálního RR intervalu, na celkovou délku 480 vzorků byly segmenty doplněny prvním způsobem doplnění (tzn. poslední hodnotou v úseku). Kromě srdečních úderů typu normální rytmus a KES byly vybírány i údery typu SVES, BLRT a BPRT). Data byla také normalizována do rozsahu od 0 do 1. Ke každému úseku také patřila anotace typu rytmu, kde 0 značila normální rytmus, 1 KES a 2 ostatní arytmie.

V tabulce 4.5 je zaznamenáno rozdělení záznamů do trénovací, validační a testovací množiny, kde N znamená počet srdečních úderů s normálním rytmem, V značí počet úderů s KES a A je počet segmentů ostatních arytmií.

Tabulka 4.5 Rozdělení dat pro klasifikaci do tří kategorií

množina	číslo záznamu	N	V	A
trénovací	100, 101, 103, 105, 106, 108, 109, 112, 113, 115, 117, 118, 121, 123, 200, 201, 202, 205, 207, 208, 212, 213, 219, 232, 233, 234	43278	3965	10186
validační	111, 124, 203, 215, 220, 222, 223, 228, 230	15701	1492	4034
testovací	114, 116, 119, 122, 209, 210, 214, 221, 231	15522	1445	3649

Celkově je tak z databáze vybráno 99272 srdečních úderů, z toho 74501 s normálním rytmem, 6902 typu KES a 17869 úderů s jinou arytmií, přesněji 2544 SVES, 8071 BLRT a 7254 BPRT. 75,05 % ze všech úseků tedy tvoří úseky normálního rytmu, 6,95 % úseky KES a 18 % ostatní arytmiie (přesněji 2,56 % SVES, 8,13 % BLRT a 7,31 % BPRT).

V tabulce 4.6 jsou počty srdečních úderů ostatních arytmií v každém záznamu přiblíženy detailněji, # značí číslo záznamu, A je počet úseků SVES, L počet segmentů BLRT a R počet BPRT; počty srdečních úderů normálního rytmu a KES jsou vypsány v tabulce 4.2.

Tabulka 4.6 Počty srdečních úderů dalších arytmií v jednotlivých záznamech

#	A	L	R	#	A	L	R	#	A	L	R
100	33	0	0	119	0	0	0	213	25	0	0
101	3	0	0	121	1	0	0	214	0	2001	0
103	2	0	0	122	0	0	0	215	3	0	0
105	0	0	0	123	0	0	0	219	7	0	0
106	0	0	0	124	2	0	1530	220	94	0	0
108	4	0	0	200	30	0	0	221	0	0	0
109	0	2491	0	201	30	0	0	222	208	0	0
111	0	2122	0	202	36	0	0	223	72	0	0
112	2	0	0	203	0	0	0	228	3	0	0
113	0	0	0	205	3	0	0	230	0	0	0
114	10	0	0	207	106	1457	85	231	1	0	1253
115	0	0	0	208	0	0	0	232	1381	0	397
116	1	0	0	209	383	0	0	233	7	0	0
117	1	0	0	210	0	0	0	234	0	0	0
118	96	0	2164	212	0	0	1825	celkem	2544	8071	7254

Trénovací množina je tvořena 57429 segmenty, z toho je 3965 (6,90 %) úderů typu KES a 10186 (17,74 %) úseků ostatních arytmií, z nich 1767 (3,08 %) úseků je SVES, 3948 (6,87 %) BLRT a 4471 (7,79 %) BPRT. Validační množina se skládá z 21227 srdečních úderů, z nich je 1492 (7,03 %) KES a 4034 (19 %) segmentů ostatních arytmií, přesněji 382 (1,80 %) segmentů SVES, 2122 (9,99 %) BLRT a 1530 (7,21 %) BPRT. Testovací množinu tvoří 20616 segmentů, z toho 1445 (7,01 %) je typu KES a 3649 (17,70 %) segmentů dalších arytmií, jmenovitě 395 (1,92 %) úseků SVES, 2001 (9,70 %) BLRT a 1253 (6,08 %) BPRT. Trénovací množina tak obsahuje 57,85 % všech dat, validační 21,38 % a testovací 20,77 %.

Ke klasifikaci srdečního rytmu do tří tříd byla použita téměř stejná architektura jako při klasifikaci do dvou tříd (viz kapitola 4.3.1 a tabulka 4.3), pouze v poslední plně propojené vrstvě byly použity 3 neurony místo jednoho a aktivační funkce softmax místo sigmoidální. Stejná architektura byla použita primárně kvůli možnosti porovnání dosažených výsledků klasifikace do dvou i tří tříd.

Při trénování probíhala optimalizace stejných hyperparametrů jako u klasifikace do dvou tříd (viz kapitola 4.3.2), jmenovitě kroku učení, počtu trénovacích epoch, velikosti dávky dat a parametru *rate* v dropout vrstvě. Hodnoty parametrů byly také vybírány ve stejném rozsahu, jelikož však změna těchto parametrů a zejména změna kroku učení v rozsahu od 0,001 do 0,0001 nevedla k velké změně, byly poté zkoušeny i hodnoty v jiném rozsahu. U kroku učení se jednalo o rozsah hodnot 0,00001 do 0,0001 s krokem 0,00001, počet epoch učení byl proto zvýšen na 40. U velikosti dávky dat byla testována i hodnota 16, a u parametru *rate* navíc hodnoty 0,45 a 0,5. Příliš malý krok učení či příliš velká *batchsize* vedla k natrénované síti, která všechny segmenty bez ohledu na skutečný rytmus klasifikovala jako normální rytmus, příliš malý *rate* vedl k naprostému přeučení sítě. V tabulce 4.7 jsou vypsány zvolené hyperparametry sítě.

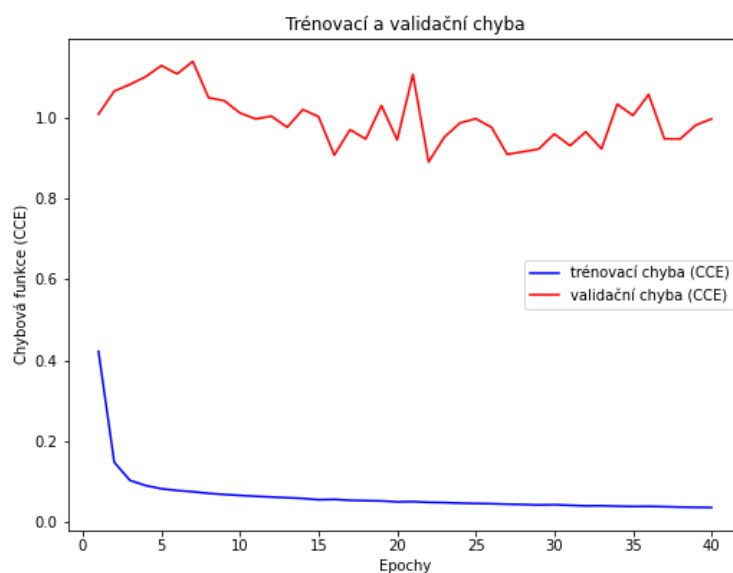
Tabulka 4.7 Nastavení hyperparametrů sítě pro klasifikaci do tří tříd

optimalizační funkce	Adam
chybová funkce	kategorická křížová entropie
metrika	přesnost
krok učení	0,00005
počet epoch	40
batchsize	64
rate (dropout)	0,4

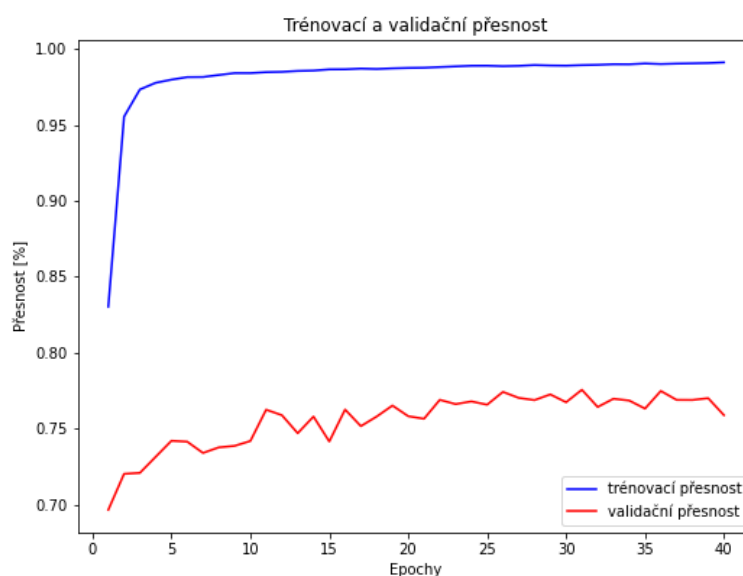
Vektor anotací obsahující hodnoty 0 až 2, které značily jednotlivé třídy, byl převeden pomocí tzv. *one-hot* kódování, kdy je každý label zakódován místo jedné skalární hodnoty do binárního vektoru o délce odpovídající počtu tříd (v tomto případě 3), kde 1 označuje příslušnost dané třídě. Segment typu KES, který byl původně anotován hodnotou 1, byl tak nově označen jako 010, a segment ostatní arytmie, původně označen

hodnotou 2, tak byl nyní anotován jako 001. *One-hot* kódování bylo použito s ohledem na ulehčení při trénování sítě v Pythonu.

Průběh trénování sítě pro klasifikaci do tří tříd je vidět na obrázcích 4.16 a 4.17. Jak lze z křivek pozorovat, již od první epochy došlo k velmi výraznému přeučení sítě. Na průběhu trénovací přesnosti i chybové funkce je vidět, že se síť naučila trénovací data klasifikovat poměrně přesně. U validačních dat sice měla přesnost v průběhu učení rostoucí tendenci, nicméně k hodnotám trénovací přesnosti se ani nepřiblížila, byla o téměř 25 % nižší. Validační chyba přibližně klesala po dobu 20 epoch a poté měla opět rostoucí tendenci.



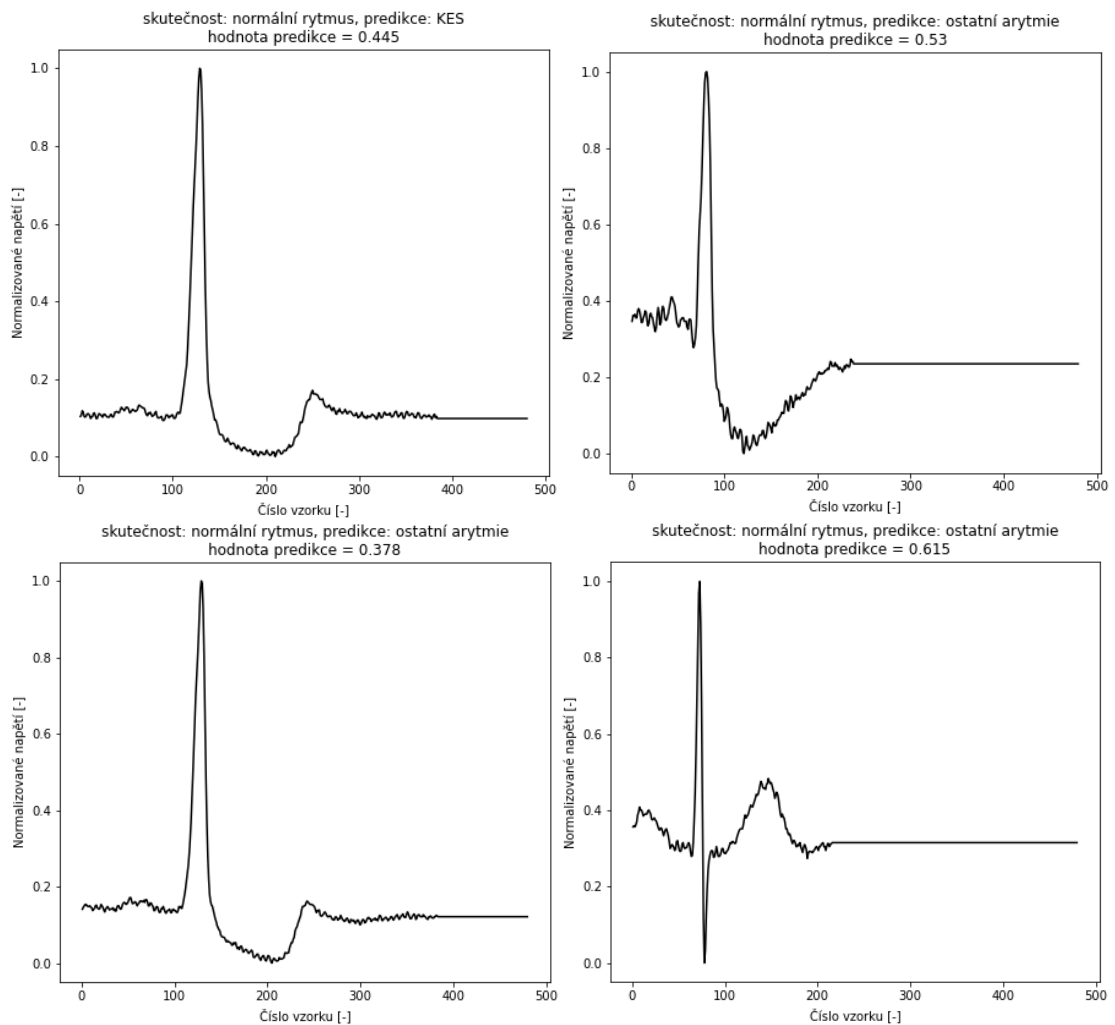
Obrázek 4.16 Průběh chybové funkce při trénování sítě pro klasifikaci do tří tříd



Obrázek 4.17 Průběh trénovací a validační přesnosti při trénování sítě pro klasifikaci do tří tříd

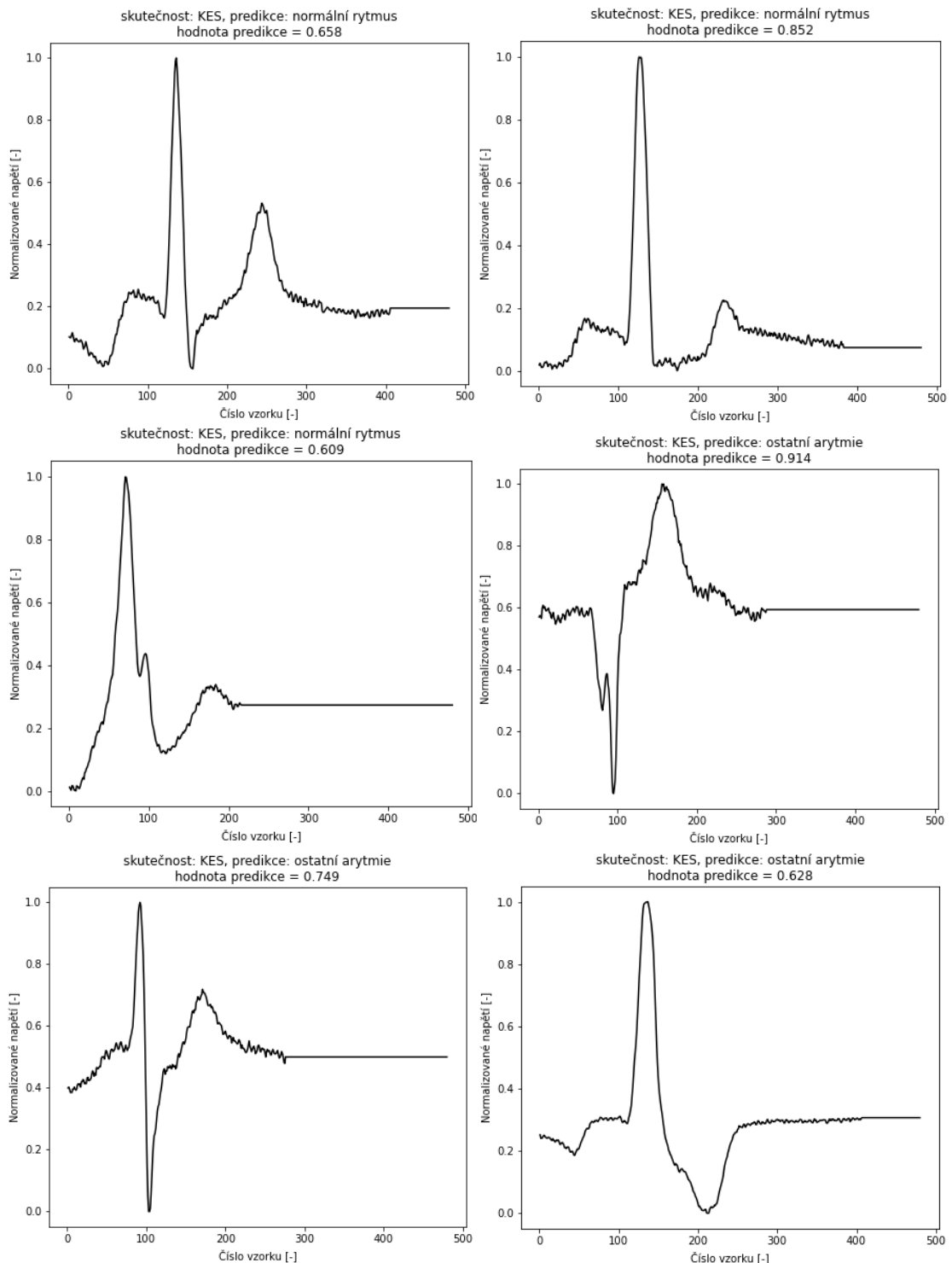
Nejvyšší přesnosti na validační množině bylo dosaženo v 31. epoše, model s vahami z této epochy byl uložen. Přesnost klasifikace pro trénovací záznamy dosáhla 98,97 %, pro validační množinu dat se rovnala 77,55 % a pro testovací 86,88 %. Velikost chybové funkce (kategorické křížové entropie) pro trénovací data vyšla 0,0385, pro validační data 0,9311 a pro testovací 0,5755. Velikost chyby pro validační data je tak více než 24krát vyšší než chyba na trénovacích datech, u testovacích dat je chyba vyšší přibližně 15krát. To je další důkaz o velkém přeučení sítě a neschopnosti správně klasifikovat segmenty EKG, které nebyly využity při učení. Výsledky jsou podrobněji shrnuty v kapitole 5.2.2, možné příčiny přeučení jsou diskutovány v kapitole 5.4.

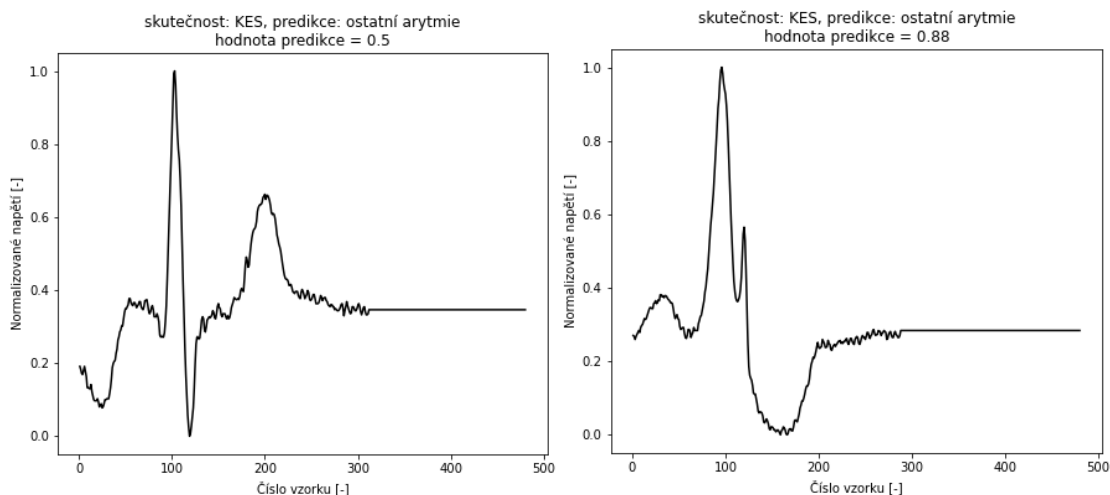
Na následujících obrázcích jsou zobrazeny příklady segmentů z testovací množiny dat, které byly klasifikovány chybně. Stejně jako u klasifikace do dvou tříd pochází všechny signály z testovací množiny dat ze svodu MLII. V titulku každého obrázku je napsána skutečná a predikovaná anotace a hodnota predikce zvolené třídy (výstup softmax funkce v poslední plně propojené vrstvě).



Obrázek 4.18 Ukázka chybně klasifikovaných segmentů normálního rytmu

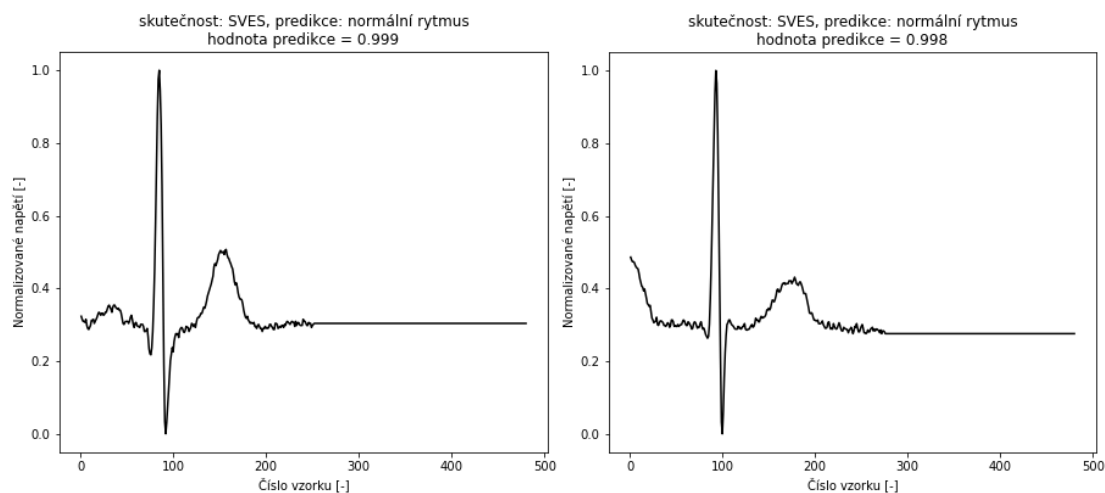
Jak je vidět na obrázku 4.18, dva téměř stejné segmenty byly klasifikovány jednou jako KES a podruhé jako ostatní arytmie, přitom hodnota predikce byla v obou případech nižší než 0,45. To znamená, že tyto segmenty nebyly do dané množiny klasifikátorem zařazeny s velkou jistotou. Všechny chybně klasifikované segmenty normálního rytmu měly obdobný průběh jako segmenty z obrázku 4.18, hodnota predikce není ani v jednom případě velká.





Obrázek 4.19 Ukázka chybně klasifikovaných segmentů KES

U chybně klasifikovaných segmentů KES je větší variabilita tvaru QRS komplexu i P a T vln, jak lze pozorovat na obrázku 4.19. Většina špatně zařazených segmentů byla klasifikována jako ostatní arytmie, často s velkou jistotou (hodnotou predikce).

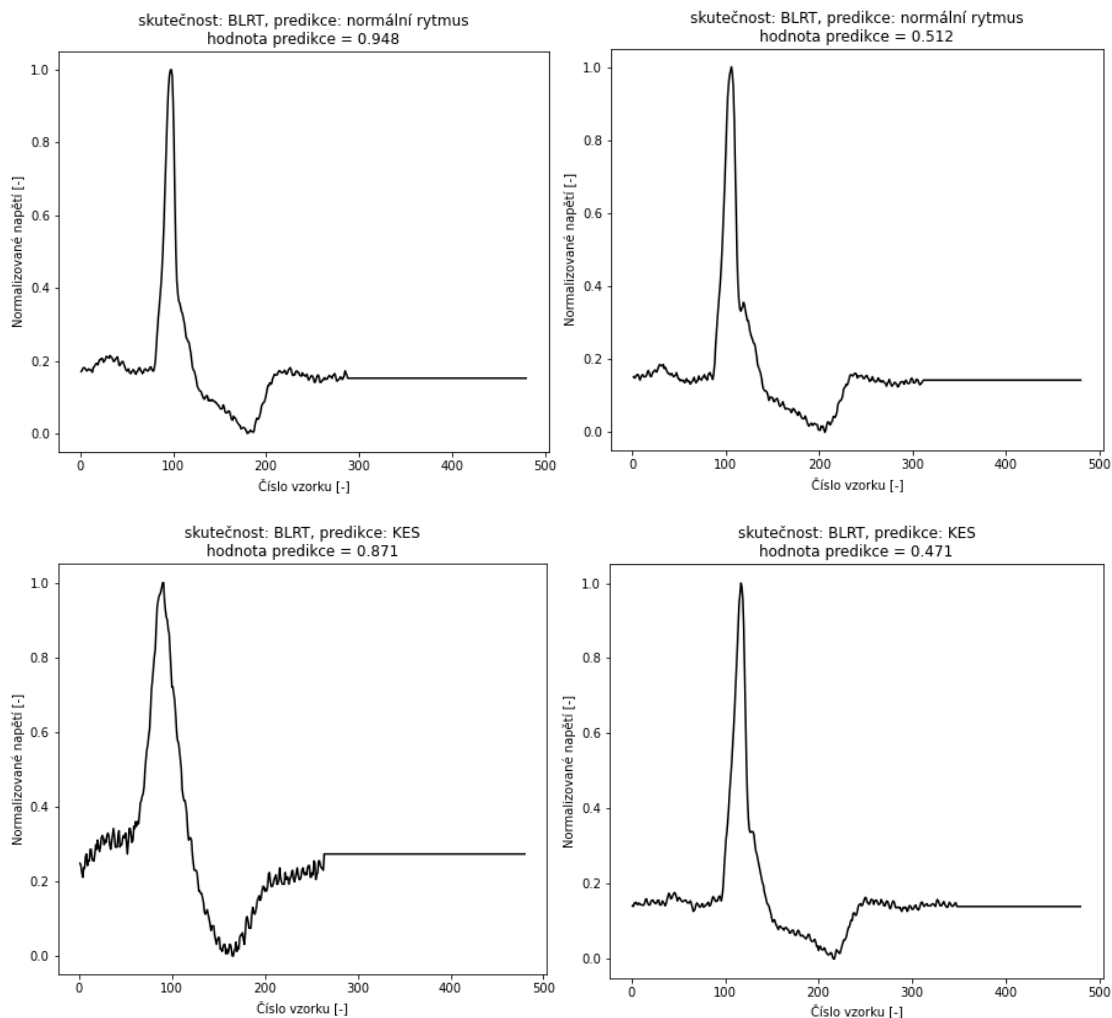


Obrázek 4.20 Ukázka chybně klasifikovaných segmentů ostatních arytmií: SVES

Prakticky všechny testovací segmenty typu SVES byly chybně klasifikovány jako normální rytmus, jako KES nebyl klasifikován ani jeden z nich. Jak je vidět z obrázku 4.20, tvar QRS komplexu normálního rytmu a SVES je stejný nebo velmi podobný, pouze vlna P může mít odlišný tvar. Hodnota predikce je tak ve většině případů rovna přibližně jedné, klasifikátor si je při zařazení těchto signálů velmi jistý.

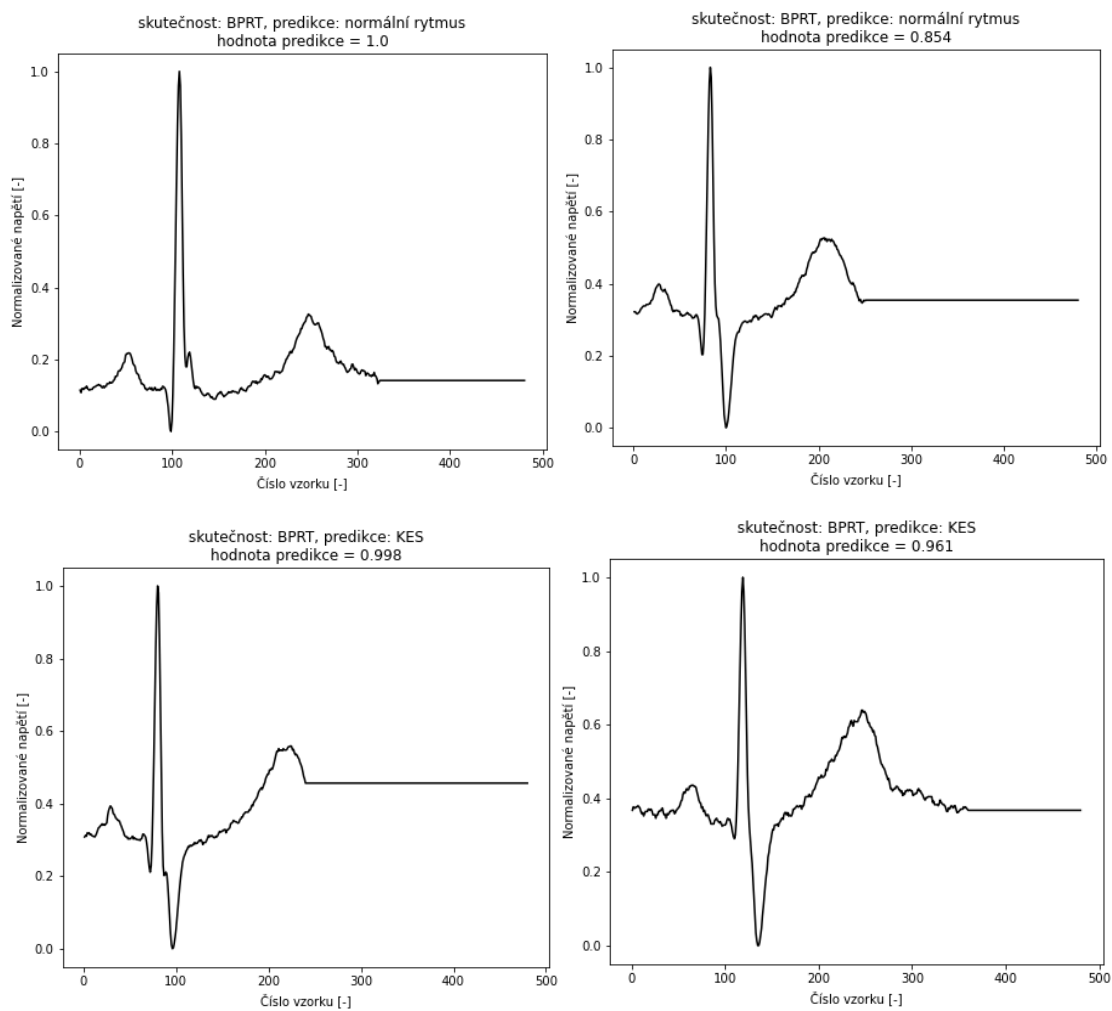
Většina úseků typu BLRT v testovací množině byla klasifikována jako normální rytmus. Na obrázku 4.21 je vidět, že téměř všechny úseky mají velmi podobný průběh, což je dané tím, že kvůli malému počtu záznamů s BLRT je v testovací množině zařazen pouze jeden záznam obsahující tuto poruchu vedení vzruchu. Některé segmenty byly

i přes obdobný tvar označeny jako KES nebo klasifikovány správně s různou hodnotou predikce.



Obrázek 4.21 Ukázka chybně klasifikovaných segmentů ostatních arytmií: BLRT

Stejně jako v případě BLRT byl kvůli nízkému počtu dostupných záznamů obsahujících BPRT v testovací množině zařazen pouze jeden záznam s tímto typem rytmu. Průběhy všech segmentů na obrázku 4.22 jsou tak opět velmi podobné. Pouze zde ukázané 2 úseky BPRT byly klasifikátorem označeny jako normální rytmus, zbylé chybně klasifikované úseky BPRT byly označeny jako KES. Hodnota predikce u většiny špatně klasifikovaných záznamů byla větší než 0,85, tyto záznamy tedy byly do dané množiny zařazeny s velkou jistotou.



Obrázek 4.22 Ukázka chybně klasifikovaných segmentů ostatních arytmií: BPRT

5. HODNOCENÍ ÚSPĚŠNOSTI KLASIFIKACE

5.1 Použité parametry pro hodnocení úspěšnosti

Úspěšnost klasifikace srdečního rytmu byla hodnocena pomocí pěti parametrů: senzitivita (Se), specificita (Sp), pozitivní prediktivní hodnota (PPV^{31}), přesnost (Acc) a F1 skóre. Senzitivita ukazuje procento správně klasifikovaných segmentů s KES, specificita naopak vyjadřuje schopnost klasifikátoru správně rozpoznat segmenty signálu s normálním rytmem. Pozitivní prediktivní hodnota říká, kolik procent ze segmentů klasifikovaných jako KES je klasifikováno správně. Přesnost je určena jako podíl správně klasifikovaných segmentů vůči všem segmentům (klasifikovaným správně i špatně). Přesnost tedy udává schopnost klasifikátoru správně zařadit segmenty podle jejich odpovídajícího rytmu. Pokud jsou však data výrazně nevyvážená a početněji zastoupená třída je klasifikována úspěšně, dosáhne tato metrika vysoké hodnoty i v případě, že je méně zastoupená třída klasifikována špatně. Výsledky klasifikace jsou pak přehnaně optimistické. F1 skóre je harmonický průměr senzitivity a pozitivní prediktivní hodnoty, tato metrika je vhodným ukazatelem při hodnocení klasifikace nevyvážených dat. Všechny parametry se nejčastěji vyjadřují v procentech. Senzitivita, specificita, PPV, přesnost a F1 skóre se vypočítají podle vzorců

$$Se = \frac{TP}{TP + FN'} \quad (5.1)$$

$$Sp = \frac{TN}{TN + FP'} \quad (5.2)$$

$$PPV = \frac{TP}{TP + FP'} \quad (5.3)$$

$$Acc = \frac{TP + TN}{TP + FP' + TN + FN'} \quad (5.4)$$

$$F1 = 2 \cdot \frac{PPV \cdot Se}{PPV + Se} \quad (5.5)$$

kde TP^{32} je počet segmentů, které jsou klasifikovány jako KES podle anotací z MIT-BIH AD i podle klasifikátoru, TN^{33} je počet segmentů, které jsou anotacemi i klasifikátorem označeny jako normální, FP^{34} je počet segmentů anotovaných databází

³¹ z anglického Positive Predictive Value

³² z anglického True Positive (pravdivě pozitivní)

³³ z anglického True Negative (pravdivě negativní)

³⁴ z anglického False Positive (falešně pozitivní)

jako normální rytmus, které klasifikátor špatně označil jako KES, a FN^{35} je počet segmentů značených podle databáze jako KES, které klasifikátor chybně klasifikoval jako normální rytmus.

5.2 Hodnocení úspěšnosti navrženého klasifikátoru

5.2.1 Klasifikace do dvou tříd

Jak již bylo psáno v kapitole 4.3.3, přesnost klasifikace na záznamech připravených prvním způsobem se pro trénovací a validační množinu rovnala 99,57 %, pro trénovací pak 99,40 %. V následujících maticích záměn jsou shrnuty počty segmentů klasifikovaných jako TP, TN, FP a FN. Tabulka 5.1 zobrazuje hodnoty pro trénovací data, tabulka 5.2 hodnoty pro validační data a tabulka 5.3 hodnoty testovacích dat. Z hodnot ze všech tří tabulek jsou následně spočítány všechny parametry pro hodnocení úspěšnosti podle vzorců definovaných v kapitole 5.1, jsou shrnuty v tabulce 5.4.

Tabulka 5.1 Matice záměn pro trénovací data doplněná prvním způsobem

trénovací data		predikovaná hodnota	
		normální rytmus	KES
skutečná hodnota	normální rytmus	51758	186
	KES	57	4768

Celkem 186 segmentů normálního rytmu bylo nesprávně klasifikováno jako KES a naopak 57 segmentů KES jako normální rytmus. Z celkového počtu 56769 trénovacích úseků jich tak bylo 243 (0,43 %) klasifikováno špatně, falešně pozitivní klasifikace byla více než třikrát častější.

Tabulka 5.2 Matice záměn pro validační data doplněná prvním způsobem

validační data		predikovaná hodnota	
		normální rytmus	KES
skutečná hodnota	normální rytmus	11327	16
	KES	37	1006

³⁵ z anglického False Negative (falešně negativní)

U validační množiny došlo naopak k více falešně negativním detekcím, kdy bylo 37 segmentů KES chybně označeno jako normální rytmus, dalších 16 úseků normálního rytmu pak bylo klasifikováno jako KES. Z celkového počtu 12386 validačních úseků jich bylo 53 (0,43 %) klasifikováno chybně.

Tabulka 5.3 Matice záměn pro testovací data doplněná prvním způsobem

testovací data		predikovaná hodnota	
		normální rytmus	KES
skutečná hodnota	normální rytmus	11194	20
	KES	53	981

U testovací množiny bylo stejně jako u validační množiny více falešně negativních detekcí. 53 úseků KES bylo sítí nesprávně klasifikováno jako normální rytmus, 20 úseků normálního rytmu bylo označeno jako KES. Z celkem 12248 testovacích segmentů jich bylo 73 (0,60 %) klasifikováno chybně.

Tabulka 5.4 Parametry hodnocení úspěšnosti klasifikace pro data doplněná prvním způsobem

množina	Se [%]	Sp [%]	PPV [%]	Acc [%]	F1 [%]
trénovací	98,82	99,64	96,25	99,57	97,52
validační	96,45	99,86	98,43	99,57	97,43
testovací	94,87	99,82	98,00	99,40	96,41

Jak je z tabulky 5.4 vidět, výsledky pro trénovací a validační množinu jsou přibližně srovnatelné, pro testovací jsou horší. Tento jev, kdy jsou výsledky pro validační množinu lepší než pro testovací, i když se síť neučila ani na jedné z těchto množin, je velmi pravděpodobně způsoben formou trénování modelu, kdy se uložily váhy sítě z té epochy, kdy bylo dosaženo nejvyšší validační přesnosti. Hyperparametry sítě byly navíc upravovány podle výsledků získaných pro validační množinu dat, dá se tedy předpokládat, že budou výsledky pro validační data o něco lepší než pro testovací.

Senzitivita klasifikátoru pro testovací množinu dosáhla pouze 94,87 %, více než 5 % segmentů s KES tedy bylo klasifikováno jako normální rytmus. Specificita je vysoká pro všechna data, jelikož došlo k malému počtu falešně pozitivních detekcí. Je to ale ovlivněno také tím, že normálních segmentů je přibližně 10krát více než segmentů s KES. PPV trénovacích dat je nižší než u validačních či testovacích, je to kvůli poměrně velkému množství falešně pozitivních detekcí. I když je celková přesnost pro trénovací i validační množinu stejná, F1 skóre je o trochu vyšší u trénovací množiny.

Pro možnost srovnání byly parametry hodnocení nakonec spočítány pro data ze všech tří množin dohromady, senzitivita dosáhla 97,87 %, specificita 99,70 %, PPV 96,82 %, přesnost 99,55 % a F1 skóre 97,34 %.

Přesnost klasifikace pro signály připravené druhým způsobem u trénovací množiny dosáhla 99,54 %, u validační 99,53 % a pro trénovací množinu se rovnala 99,41 %. Tabulka 5.5 představuje matici záměn pro trénovací data, tabulka 5.6 matici záměn pro validační data a tabulka 5.7 matici záměn testovacích dat.

Tabulka 5.5 Matice záměn pro trénovací data doplněná druhým způsobem

trénovací data		predikovaná hodnota	
		normální rytmus	KES
skutečná hodnota	normální rytmus	51778	166
	KES	93	4732

Dohromady 166 úseků s normálním rytmem bylo chybně klasifikováno jako KES a naopak 93 úseků s KES jako normální rytmus. Z celkem 56769 trénovacích úseků jich bylo 259 (0,46 %) klasifikováno nesprávně, falešně pozitivní klasifikace zhruba dvakrát častější.

Tabulka 5.6 Matice záměn pro validační data doplněná druhým způsobem

validační data		predikovaná hodnota	
		normální rytmus	KES
skutečná hodnota	normální rytmus	11335	8
	KES	50	993

U validační množiny došlo oproti falešně pozitivním klasifikacím k více než pětinasobku falešně negativních klasifikací. 50 segmentů KES bylo nesprávně klasifikováno jako normální rytmus, 8 úseků normálního rytmu bylo označeno jako KES. Celkem 58 (0,47 %) z 12386 validačních úseků tak bylo klasifikováno chybně.

U testovací množiny bylo falešně negativních detekcí dokonce osmkrát více než falešně pozitivních. Celkově 64 úseků KES bylo sítí nesprávně označeno jako normální, 8 segmentů normálního rytmu bylo klasifikováno jako KES. Z celkového počtu 12248 testovacích segmentů jich tak bylo 72 (0,59 %) klasifikováno chybně.

Tabulka 5.7 Matice záměn pro testovací data doplněná druhým způsobem

testovací data		predikovaná hodnota	
		normální rytmus	KES
skutečná hodnota	normální rytmus	11206	8
	KES	64	970

V tabulce 5.8 jsou shrnuty parametry pro hodnocení úspěšnosti pro data připravená druhým způsobem doplnění.

Tabulka 5.8 Parametry hodnocení úspěšnosti klasifikace pro data doplněná druhým způsobem

množina	Se [%]	Sp [%]	PPV [%]	Acc [%]	F1 [%]
trénovací	98,07	99,68	96,61	99,54	97,34
validační	95,20	99,93	99,20	99,53	97,16
testovací	93,81	99,93	99,18	99,41	96,42

Výsledky dosažené na trénovací a validační množině jsou opět přibližně srovnatelné, pro testovací množinu jsou o něco horší. Senzitivita klasifikace pro testovací množinu dosáhla jen 93,81 %, 64 segmentů KES bylo totiž klasifikováno jako normální rytmus. I senzitivita pro validační množinu je nižší, opět je to způsobeno velkým množstvím falešně negativních detekcí. Specificita je vysoká pro všechny množiny dat, pro testovací i validační je dokonce vyšší než pro trénovací. PPV trénovacích dat je kvůli velkému počtu falešně pozitivních detekcí nižší než u obou zbývajících datových množin. Přesnost je vysoká u všech množin, F1 skóre je i přes nižší PPV nejvyšší u trénovací množiny. Senzitivita modelu pro všechna data dohromady dosáhla 97,00 %, specificita 99,76 %, PPV 97,35 %, přesnost 99,52 % a F1 skóre 97,18 %.

Při srovnání celkových výsledků dosažených pro oba způsoby na všech (trénovacích, validačních i testovacích) záznamech současně (viz tabulka 5.9) lze pozorovat, že jsou dosažené výsledky podobné. U prvního způsobu doplnění bylo oproti druhému způsobu o 60 FN detekcí méně, díky tomu bylo pro tento způsob dosaženo vyšší senzitivity. U druhého způsobu došlo naopak k méně FP klasifikacím, proto je specificita o 0,06 % vyšší a PPV je dokonce vyšší o přibližně 0,5 %. Přesnost pro první způsob doplnění je nepatrně vyšší, celkový počet chybně klasifikovaných úseků je 369 z 81403, což odpovídá 0,45 %. U druhého způsobu došlo k celkem 389 nesprávným detekcím, to odpovídá 0,48 % chybně klasifikovaným vzorkům. F1 skóre je o 0,16 % vyšší pro první způsob doplnění vzorků, první způsob doplnění tedy byl mírně úspěšnější.

Tabulka 5.9 Porovnání úspěšnosti klasifikace pro oba způsoby doplnění vzorků

způsob doplnění	TP	TN	FP	FN	Se [%]	Sp [%]	PPV [%]	Acc [%]	F1 [%]
první	6755	75279	222	147	97,87	99,70	96,82	99,55	97,34
druhý	6695	74319	182	207	97,00	99,76	97,35	99,52	97,18

S ohledem na tyto výsledky lze celkově konstatovat, že síť dosáhla pro oba způsoby doplnění přibližně stejných výsledků a způsob doplnění úseků na stejnou délku tak nehraje příliš významnou roli.

5.2.2 Klasifikace do tří tříd

Jak bylo psáno v kapitole 4.4, u klasifikace do tří tříd došlo k výraznému přeučení sítě. Přesnost klasifikace se pro trénovací záznamy rovnala 98,97 %, pro validační data vyšla 77,55 % a pro testovací 86,88 %. V následujících maticích záměn jsou shrnuty počty správně a špatně klasifikovaných segmentů, úseky na diagonále těchto matic jsou klasifikovány správně, mimo diagonálu chybně. Pro každou množinu dat je zároveň vypsána tabulka s celkovými počty TP, TN, FP a FN klasifikací a parametry pro hodnocení úspěšnosti. Tabulky 5.10 a 5.11 zobrazují výsledky pro trénovací data, tabulky 5.12 a 5.13 hodnoty pro validační data a tabulky 5.14 a 5.15 výsledky testovacích dat.

Tabulka 5.10 Matice záměn pro trénovací data

trénovací data		predikovaná hodnota		
		normální rytmus	KES	ostatní arytmie
skutečná hodnota	normální rytmus	43117	36	125
	KES	37	3837	91
	ostatní arytmie	169	132	9885

36 segmentů normálního rytmu bylo klasifikováno jako KES a 125 jako ostatní arytmie. Z celkem 43278 úseků normálního rytmu jich tak 161 (0,37 %) bylo klasifikováno špatně. Při klasifikaci KES bylo nesprávně zařazeno 128 úseků z 3965 (3,23 %), z toho 37 jako normální rytmus a 91 jako ostatní arytmie. K nejvíce chybám došlo u klasifikace úseků ostatních arytmií, špatně bylo zařazeno celkem 301 úseků z 10186 (2,96 %). Jako KES bylo klasifikováno 132 segmentů ostatních arytmií, z toho 48 SVES, 43 BLRT a 41 BPRT. 169 úseků pak bylo chybně označeno jako normální rytmus, z toho 158 SVES, 5 BLRT a 6 BPRT. Z celkového počtu 1767 SVES jich tak 206 (11,66 %) bylo zařazeno chybně, z 3948 BLRT pak bylo nesprávně označeno 48 úseků (1,22 %) a z 4471 BPRT

úseků jich bylo chybně klasifikováno 47 (1,05 %). Z celkového počtu 57429 trénovacích úseků jich tak bylo 590 (1,03 %) klasifikováno špatně, u normálního rytmu i KES bylo více úseků klasifikováno jako ostatní arytmie, chybně zařazené úseky ostatních arytmií pak byly častěji označeny jako normální rytmus.

Tabulka 5.11 Úspěšnost klasifikace trénovacích dat pro všechny tři třídy

třída	TP	TN	FP	FN	Se [%]	Sp [%]	PPV [%]	Acc [%]	F1 [%]
normální rytmus	43117	13945	206	161	99,63	98,54	99,52	99,36	99,58
KES	3837	53296	168	128	96,77	99,69	95,81	99,48	96,29
ostatní arytmie	9885	47027	216	301	97,04	99,54	97,86	99,10	97,45

Z tabulek 5.10 i 5.11 lze vyčíst, že síť byla schopna trénovací data klasifikovat s vysokou přesností. U normálního rytmu a KES bylo více FP klasifikací, u ostatních arytmií naopak FN. Senzitivita, PPV i z nich odvozené F1 skóre je nejvyšší u normálního rytmu díky velkému počtu TP segmentů, naopak specificita je díky velkému množství TN klasifikací vyšší u KES i ostatních arytmií. Nejnižší senzitivita, PPV i F1 skóre vyšlo u třídy KES.

Tabulka 5.12 Matice záměn pro validační data

validační data		predikovaná hodnota		
		normální rytmus	KES	ostatní arytmie
skutečná hodnota	normální rytmus	13923	563	1215
	KES	146	920	426
	ostatní arytmie	2218	197	1619

Z tabulky 5.12 je kvůli velkému počtu nesprávných klasifikací na první pohled vidět, že byla síť přeučena. Z celkového počtu 15701 úseků normálního rytmu jich bylo 1778 (11,32 %) zařazeno špatně, z toho více než 2/3 do třídy ostatní arytmie. U třídy KES bylo chybně klasifikováno 572 segmentů z 1492 (38,34 %), více segmentů bylo opět zařazeno do třídy ostatní arytmie. Nejhůře dopadla klasifikace pro třídu ostatní arytmie, celkem 2415 úseků z 4034 (59,87 %) bylo predikováno chybně. Jako KES bylo označeno 197 segmentů ostatních arytmií, všechny takto zařazené úseky byly typu BLRT. 2218 úseků bylo nesprávně klasifikováno jako normální rytmus, z toho 378 SVES, 348 BLRT a 1492 BPRT. Z dohromady 382 SVES úseků tak byly pouze 4 zařazeny správně. Z celkem 2122 BLRT bylo nesprávně zařazeno 545 úseků (25,68 %), chybně bylo klasifikováno také

více než 97 % úseků BPRT z celkového počtu 1530. Dohromady 4765 validačních úseků (22,45 %) z 21227 bylo klasifikováno chybně, u normálního rytmu i KES bylo více úseků klasifikováno jako ostatní arytmie, segmenty ostatních arytmií byly nejčastěji označeny jako normální rytmus.

Tabulka 5.13 Úspěšnost klasifikace validačních dat pro všechny tři třídy

třída	TP	TN	FP	FN	Se [%]	Sp [%]	PPV [%]	Acc [%]	F1 [%]
normální rytmus	13923	3162	2364	1778	88,68	57,22	85,49	80,49	87,05
KES	920	18975	760	572	61,66	96,15	54,76	93,72	58,01
ostatní arytmie	1619	15552	1641	2415	40,13	90,46	49,66	80,89	44,39

Klasifikace nejlépe dopadla pro úseky normálního rytmu, senzitivita, PPV i F1 skóre je výrazně vyšší než pro KES a ostatní arytmie. Nízká specifická se dá vysvětlit velkým počtem FN segmentů. Naopak vysoká senzitivita u zbývajících tříd (především u KES) je způsobena velkým množstvím TN segmentů. Nejhorší dopadla klasifikace pro třídu ostatních arytmií, kdy byla senzitivita i skóre F1 nižší než 45 % a přibližně 60 % úseků z této třídy bylo zařazeno špatně.

Tabulka 5.14 Matice záměn pro testovací data

testovací data		predikovaná hodnota		
		normální rytmus	KES	ostatní arytmie
skutečná hodnota	normální rytmus	15393	15	114
	KES	61	1049	335
	ostatní arytmie	2122	57	1470

Z tabulky 5.14 se opět dá poznat přeučení sítě, především pro třídu ostatních arytmií. 129 segmentů (0,83 %) normálního rytmu z 15522 bylo klasifikováno chybně, téměř 8krát častější bylo označení úseku jako ostatní arytmie. Při klasifikaci KES bylo nesprávně zařazeno 396 segmentů z 1445 (27,40 %), většina chyb plynula z označení segmentů jako ostatní arytmie. Stejně jako u validační množiny dopadla klasifikace nejhorší pro třídu ostatní arytmie, špatně bylo klasifikováno celkem 2179 úseků z 3649 (59,71 %). Jako KES bylo klasifikováno 57 segmentů třídy ostatní arytmie, přesněji 1 SVES, 3 BLRT a 53 BPRT. 2122 úseků pak bylo chybně označeno jako normální rytmus, z toho 389 SVES, 1731 BLRT a 2 BPRT. Celkem 390 úseků SVES z 395 bylo zařazeno špatně, a ve validační a testovací množině dohromady tak bylo pouze 9 úseků

SVES ze 777 zařazeno správně. Z celkového počtu 2001 segmentů BLRT jich bylo správně klasifikováno pouze 267 (13,34 %). Dohromady 175 úseků BPRT z 1253 (13,97 %) bylo predikováno špatně. Z celkového počtu 20616 testovacích úseků jich bylo 2704 (13,12 %) klasifikováno nesprávně, u normálního rytmu i KES převládalo chybné zařazení do třídy ostatní arytmie, nesprávně klasifikované úseky ostatních arytmií pak byly ve většině případů označeny jako normální rytmus.

Tabulka 5.15 Úspěšnost klasifikace testovacích dat pro všechny tři třídy

třída	TP	TN	FP	FN	Se [%]	Sp [%]	PPV [%]	Acc [%]	F1 [%]
normální rytmus	15393	2911	2183	129	99,17	57,15	87,58	88,79	93,01
KES	1049	19099	72	396	72,60	99,62	93,58	97,73	81,76
ostatní arytmie	1470	16518	449	2179	40,29	97,35	76,60	87,25	52,80

Při porovnání tabulek 5.13 a 5.15 lze konstatovat, že klasifikace byla úspěšnější pro záznamy z testovací množiny, především pro třídu KES. Stejně jako u validačních dat dopadla klasifikace nejlépe pro třídu normálního rytmu. U třídy KES je oproti validační množině přibližně 10krát méně FP detekcí, výrazně ubylo také FN. Mírné zlepšení se dá pozorovat i u třídy ostatních arytmií zejména díky menšímu počtu FP segmentů, nicméně senzitivita i F1 skóre je pořád velmi nízké. Zajímavé je, že u validační množiny byly nejčastějším zdrojem FN klasifikací chybně zařazené segmenty BPRT, u testovací naopak BLRT.

S ohledem na výše popsané výsledky se dá říct, že síť byla výrazně přeučena, jelikož dobrých výsledků klasifikace dosáhla pouze u trénovací množiny dat. I když síť nebyla trénována na validačních ani testovacích datech, výsledky pro obě množiny se poměrně lišily, zejména v počtu FP a FN klasifikací. Možné příčiny přeučení jsou diskutovány v kapitole 5.4.

5.3 Porovnání dosažených výsledků s publikovanými metodami

V tabulce 5.16 jsou porovnány výsledky klasifikace dosažené v rámci této práce s výsledky klasifikačních metod publikovaných v literatuře, které byly detailněji popsány v kapitole 2. Výsledky v tabulce jsou však spíše orientační a nedají se přesně porovnat, protože byly jednotlivé algoritmy testovány pokaždé na jiném rozložení dat a cílem různých metod byla klasifikace do různých tříd a kategorií. V tabulce jsou zaznačeny pouze výsledky dosažené na záznamech z MIT-BIH AD, v dalším textu jsou pak výsledky rozebrány podrobněji. # označuje počet tříd, do kterých daná metoda srdeční rytmus klasifikuje.

Tabulka 5.16 Porovnání dosažených výsledků s publikovanými metodami

Název práce	#	Se [%]	Sp [%]	PPV [%]	Acc [%]
Detection of Premature Ventricular Contraction for real-time applications [21]	2	91,05	99,55	-	-
Rule-based rough-refined two-step-procedure for real-time premature beat detection in single-lead ECG [22]	3	76,54	-	90,47	85,65
Cardiac arrhythmia classification using statistical and mixture modeling features of ECG signals [23]	16	99,70	-	99,69	-
	5	96,16	-	96,15	-
Heartbeat Classification Using Morphological and Dynamic Features of ECG Signals [24]	16	98,27	-	97,20	99,32
	5	84,70	-	60,30	88,20
Application of principal component analysis to ECG signals for automated diagnosis of cardiac health [25]	5	99,90	99,10	99,61	98,11
Effective and efficient detection of premature ventricular contractions based on variation of principal directions [26]	2	96,12	-	86,48	98,77
Automated detection of premature ventricular contraction in ECG signals using enhanced template matching algorithm [27]	2	92,01	99,24	-	98,84
A Construction Method of Personalized ECG Template and Its Application in Premature Ventricular Contraction Recognition for ECG Mobile Phones [28]	2	99,17	99,43	-	-
Premature Ventricular Contraction Beat Detection with Deep Neural Networks [29]	2	96,08	-	-	99,41
A deep learning approach for ECG-based heartbeat classification for arrhythmia detection [30]	2	98,55	99,52	-	99,09
A novel technique for cardiac arrhythmia classification using spectral correlation and support vector machines [31]	5	99,20	99,70	99,90	98,60
PVC Detection Using a Convolutional Autoencoder and Random Forest Classifier [33]	2	85,65	98,90	-	98,43
Arrhythmia Detection from 2-lead ECG using Convolutional Denoising Autoencoders [34]	5	84,70	96,60	74,80	95,70
A novel electrocardiogram arrhythmia classification method based on stacked sparse auto-encoders and softmax regression [35]	6	99,22	-	99,37	-
A deep convolutional neural network model to classify heartbeats [36]	5	96,71	91,54	97,86	94,03
A Convolutional Neural Network for Identifying Premature Ventricular Contraction Beat and Right Bundle Branch Block Beat [37]	3	98,37	99,19	-	98,92
Classification of ECG signals based on 1D convolution neural network [38]	5	-	-	-	97,50
Tato práce	2	94,87	99,82	98,00	99,40
	3	72,60	99,62	93,58	97,73

Metoda popsaná v článku [21] klasifikuje srdeční úderý na normální nebo KES na základě rozhodovacích pravidel. Byla testována na 26 záznamech z MIT-BIH AD, které obsahovaly celkem 60165 srdečních úderů, z toho 4670 KES. Senzitivita metody dosáhla 91,05 % a specificita 99,55 %.

Metoda zmíněná ve studii [22] také využívá rozhodovacích pravidel ke klasifikaci srdečních úderů na normální rytmus, KES a SVES. Algoritmus byl trénován na 1832 desetivteřinových EKG záznamech z databáze CPSC2018 (China Physiological Signal Challenge), úspěšnost algoritmu byla poté testována na signálech ze tří různých databází, a to MIT-BIH AD, Wearable ECG database a záznamy ze stejné databáze s přidáním šumem z Noise stress test database. Pro 44 záznamů z MIT-BIH AD bylo dosaženo celkové senzitivity 98,31 % pro úseky normálního rytmu, 19,10 % pro SVES a 76,54 % pro KES, PPV postupně 78,44 %, 31,79 % a 90,47 %. Celkové F1 skóre se rovnalo 64,68 % a přesnost 85,56 %. Pro záznamy z databáze Wearable ECG, pro které je algoritmus zamýšlen, se senzitivita pro normální úseky rovnala 99,70 %, pro SVES 77,06 % a KES 91,13 %, PPV postupně 97,53 %, 95,29 % a 98,17 %, celkové F1 skóre vyšlo 92,78 % a přesnost 97,51 %.

Metoda [23] využívající časových, morfologických a statistických příznaků byla testována dvěma různými způsoby rozdělení záznamů z MIT-BIH AD. U prvního způsobu (tzv. *class-oriented* přístup) jsou jednotlivé segmenty srdečních úderů ze všech záznamů náhodně přiřazeny do trénovací či testovací množiny, je tedy pravděpodobné, že se srdeční úderý z jednoho patientského záznamu objeví v trénovací i testovací množině. Tím může dojít k přeučení a výsledky klasifikace budou přehnaně optimistické. U druhého způsobu (tzv. *subject-oriented* přístup), který je více realistický, je pak celý EKG záznam od jednoho pacienta (respektive všechny segmenty) zařazen buď do trénovací nebo testovací množiny. Prvním způsobem byly jednotlivé segmenty EKG signálu klasifikovány do 16 tříd. Pro KES dosáhla klasifikace 100 % senzitivity i pozitivní prediktivní hodnoty, pro všechny třídy pak průměrné senzitivity 99,70 % a pozitivní prediktivní hodnoty 99,69 %. Druhým způsobem pak byly segmenty klasifikovány do pěti tříd, a to N (srdeční úderý, které mají původ v sinusovém uzlu), S (supraventrikulární extrasystoly), V (komorové extrasystoly), F (fúzní rytmus) a Q (neklasifikovatelné úderý). U klasifikace KES se senzitivita rovnala 95,99 % a PPV 77,63 %, celková průměrná senzitivita pak byla 96,16 % a PPV 96,15 %.

Přístup [24] používá ke klasifikaci morfologické a dynamické příznaky. Ke trénování i testování bylo použito 44 signálů z obou svodů záznamů z MIT-BIH AD. Při rozdělení dat pomocí *class-oriented* přístupu byl srdeční rytmus klasifikován do 16 tříd, pro KES bylo dosaženo senzitivity 98,27 % a PPV 97,20 %, celková přesnost se pak rovnala 99,32 %. Když byla polovina záznamů použita pro trénování a zbytek pro testování, segmenty byly klasifikovány do pěti tříd (stejně jako u předchozí zmíněné metody). Senzitivita pro KES se pak rovnala 84,7 % a pro SVES 56,4 %, PPV postupně 60,3 % a 55,2 %, celková přesnost dosáhla 88,2 %.

Další zmíněná metoda [25] využívající PCA a SVM byla použita ke klasifikaci srdečních úderů na normální rytmus, BPRT, BLRT, KES a SVES. Data byla rozdělena *class-oriented* způsobem, ke trénování a testování bylo využito 34989 srdečních cyklů z MIT-BIH AD, z toho 10000 normálních, 7250 BPRT, 8069 BLRT, 2544 SVES a 7126 KES úderů. K testování byla použita 10násobná křížová validace, bylo přitom dosaženo průměrné senzitivity 99,90 %, specifinity 99,10 %, PPV 99,61 % a přesnosti 98,11 %.

Metoda [26] k detekci KES využívá směr největšího rozptylu v datech vypočítaný z PCA. Algoritmus byl trénován na 22 záznamech z MIT-BIH AD, úspěšnost klasifikace byla testována na zbylých 22 záznamech obsahujících 44247 normálních úseků a 3220 úseků KES. Průměrná přesnost se rovnala 98,77 %, senzitivita 96,12 % a pozitivní prediktivní hodnota 86,48 %.

Podstatou metody [27] je klasifikace srdečního rytmu na normální nebo KES na základě vyhledávání vzorů mezi časovými a morfologickými příznaky. Metoda byla trénována na 12 záznamech z MIT-BIH AD, její úspěšnost byla poté testována na záznamech z pěti různých databází (MIT-BIH AD, St. Petersburg INCART 12-lead arrhythmia database, QT database, Fantasia database a MIT-BIH Supraventricular Arrhythmia database). Na všech záznamech dohromady bylo dosaženo celkové senzitivity 90,81 %, specifinity 99,40 % a přesnosti 99,00 %. Pro zbylých 36 záznamů z MIT-BIH AD, které nebyly využity v trénovací fázi, se výsledná senzitivita rovnala 92,01 %, specifita 99,24 % a přesnost 98,84 %, pro 12 vybraných záznamů ze stejné databáze pak senzitivita vyšla 96,92 %, specifita 99,81 % a přesnost 99,53 %.

Základem metody [28] využití k detekci KES je také vyhledávání vzorů. V článku byly použity záznamy z MIT-BIH AD, 21 jich sloužilo k vytvoření algoritmu a zbylých 22 k jeho otestování. Přitom byl vždy z prvních 5 minut záznamu vytvořen individualizovaný vzor, detekce KES je poté provedena na zbývajících 25 minutách záznamu. U testovacích dat bylo dosaženo senzitivity 99,17 % a specifinity 99,43 %.

V metodě [29] bylo z EKG záznamů vypočteno několik parametrů, ke klasifikaci srdečního rytmu na normální nebo KES byla poté využita DNN. Algoritmus byl natrénován a otestován pomocí 8násobné křížové validace na celkem 80636 úsecích srdečních úderů ze signálů z MIT-BIH AD, z toho 74157 úseků mělo normální rytmus a 6678 KES. Data tak byla rozdělena *class-oriented* způsobem, přitom bylo dosaženo 96,08 % senzitivity a 99,41 % celkové přesnosti.

DNN byla pro klasifikaci srdečních úderů na normální a abnormální využita i v článku [30]. Ke trénování a testování bylo využito 2288 úseků normálního rytmu, které byly náhodně vybrány z 66750 možných normálních úseků, a 2288 úseků s KES, SVES nebo fúzním rytmem. Nízký počet normálních úseků byl zvolen proto, aby byl vyvážený počet úseků z obou možných klasifikačních tříd a DNN tak neměla tendenci považovat za normální rytmus všech úseků. Trénovací sadu tvořilo 2712 segmentů, z toho 1466 normálních a 1246 abnormálních, testovací pak 1864 úseků, z toho 822 normálních

a 1042 abnormálních. Pro testovací sadu bylo dosaženo 98,55 % senzitivity, 99,52 % specificity a 99,09 % přesnosti.

Metoda prezentovaná v článku [31] využila ke klasifikaci pěti typů rytmu (normální, BPRT, BLRT, KES a SVES) spektrální korelaci a nelineární SVM. Navržený algoritmus byl otestován na 6259 srdečních úderech (z toho 1500 normálních, 1317 SVES, 1274 KES, 1131 BLRT a 1037 BPRT) z celkem 30 záznamů z MIT-BIH AD s využitím 10násobné křížové validace. Senzitivita dosáhla 99,2 %, specificita 99,7 %, PPV 99,9 % a přesnost 98,6 %. Protože však bylo pro trénování použito 17940 příznaků, v článku byla otestována i metoda s výrazně redukováným počtem pouze nejdůležitějších 88 příznaků. Úspěšnost klasifikace byla v tomto případě nižší, senzitivita i specificita se rovnala 98,9 %, PPV 99,6 % a přesnost 96,3 %.

Metoda publikovaná v článku [33] k detekci KES využívá CAE a náhodný les. Algoritmus byl testován pro tři různé výběry záznamů z MIT-BIH AD. Nejprve byla polovina úderů z každého záznamu vybrána do trénovací množiny a zbytek byl použit pro testování. Trénovací sada tak byla složena z 54695 srdečních úderů, z toho 3495 KES a testovací sada z 54675 úderů a z toho 3633 KES, toto rozdělení tedy bylo *class-oriented*. Pro tuto verzi bylo dosaženo 98,43 % přesnosti, 85,65 % senzitivity a 98,90 % specificity. Následně byla simulována nepřesná detekce R peaků tak, že byly pozice R vln posunuty náhodně až o 36 vzorků, což odpovídá detekční chybě 100 ms. Přesnost klasifikace pro takto upravená data dosáhla 97,60 %, senzitivita 84,93 % a specificita 98,42 %. Nakonec byly pro posouzení schopnosti zobecnění na nové pacienty použity pro trénování pouze srdeční úderů ze dvou záznamů a metoda byla otestována na čtyřech jiných záznamech. Přesnost se v tomto případě rovnala 87,80 %, senzitivita 86,56 % a specificita 88,09 %.

Metoda [34] ke klasifikaci srdečního rytmu do pěti tříd (N, V, S, F, Q) využívá CDAE. Algoritmus byl trénován a testován metodou pětínásobné křížové validace za použití 44 záznamů z MIT-BIH AD, kdy vždy čtyři pětiny záznamů sloužily ke trénování a zbylé záznamy k testování. Pro srdeční úderů typu KES se přesnost klasifikace rovnala 95,7 %, senzitivita 84,7 %, specificita 96,6 % a PPV 74,8 %, pro SVES pak přesnost dosáhla 96,6 %, senzitivita pouze 4,8 %, specificita 99,3 % a PPV 21,4 %.

Metoda popsána v článku [35] provádí klasifikaci šesti typů rytmu (normální, BPRT, BLRT, KES, SVES a rytmus udávaný kardiostimulátorem) pomocí čtyř autoenkodérů poskládaných za sebe. K natrénování a otestování úspěšnosti metody bylo použito 23 záznamů z MIT-BIH AD. Data byla rozdělena *class-oriented* způsobem, některé segmenty z každého záznamu byly náhodně vybrány pro trénování sítě a zbytek pro testování. Dohromady bylo použito 23200 srdečních úderů, z toho 7200 normálních, 4800 BLRT, 4800 BPRT, 2300 KES, 1700 SVES a 2400 úderů daných kardiostimulátorem. Celková senzitivita klasifikace dosáhla 99,22 % a PPV 99,37 %, senzitivita detekce KES se rovnala 98,43 % a PPV 98,18 %.

Následující tři metody ke klasifikaci srdečního rytmu využívají CNN, liší se v architektuře navržených sítí a nastavených hyperparametrech. V metodě [36] bylo ke

klasifikaci úseků do pěti tříd (N, V, S, F, Q) použito celkem 109449 srdečních úderů ze 48 záznamů z MIT-BIH AD. Jelikož je v záznamech nevyvážený počet srdečních úderů náležících do každé třídy, byly změnou směrodatné odchylky a průměru uměle vytvořeny další data tak, aby se počet segmentů každé třídy rovnal počtu úseků s normálním rytmem. Celkem tak bylo pro učení a testování k dispozici 452 960 segmentů, z toho vždy 90592 úseků náleželo do stejné třídy. Segmenty byly poté rozděleny *class-oriented* způsobem, trénování i testování modelu proběhlo pomocí 10násobné křížové validace. Celková přesnost klasifikace se rovnala 94,03 %, senzitivita 96,71 %, specificita 91,54 % a PPV 97,86 %, pouze pro KES pak byla přesnost 97,91 %, senzitivita 94,22 %, specificita 98,84 % a PPV 95,30 %.

V článku [37] byla CNN využita ke klasifikaci KES, BPRT a normálního srdečního rytmu. Algoritmus byl trénován a testován 10násobnou křížovou validací na 45 záznamech z MIT-BIH AD, čtyři nebo pět vybraných záznamů vždy sloužilo k testování, zbylé záznamy ke trénování. Celkem bylo použito 87223 srdečních úderů, z toho 72896 s normálním rytmem, 7080 s KES a 7247 s BPRT, přitom bylo dosaženo senzitivity 98,37 %, specificity 99,19 % a přesnosti 98,92 %. Pouze pro KES se pak senzitivita rovnala 88,97 %, specificita 99,98 % a přesnost 98,46 %.

Metoda představená v článku [38] CNN využila ke klasifikaci signálů na normální, KES, SVES, BPRT a BLRT. Data byla rozdělena *class-oriented* přístupem, neuronová síť byla trénována na 13200 náhodně vybraných srdečních úderech z MIT-BIH AD a testována na stejně velkém počtu jiných náhodně vybraných úderů, přitom bylo dosaženo přesnosti 97,5 %.

V této práci byla ke klasifikaci záznamů do dvou i tří tříd využita CNN s LSTM jednotkou. Při klasifikaci na normální rytmus a KES byla síť trénována na 32 záznamech z MIT-BIH AD, dalších 5 záznamů sloužilo k validaci a 6 k testování, data tak byla rozdělena pomocí *subject-oriented* přístupu. Data byla připravena dvěma způsoby, při prvním bylo na testovací skupině dat dosaženo senzitivity 94,87 %, specificity 99,82 %, PPV 98,00 %, přesnosti 99,40 % a skóre F1 96,41 %. U druhého způsobu doplnění se senzitivita u testovací množiny dat rovnala 93,81 %, specificita 99,93 %, PPV 99,18 %, přesnost 99,41 % a skóre F1 96,42 %. Při klasifikaci na normální rytmus, KES a ostatní arytmie (SVES, BLRT a BPRT dohromady) bylo ke trénování použito 26 záznamů, k validaci i testování vždy 9. Data byla i v tomto případě rozdělena *subject-oriented* přístupem. U testovací množiny dat bylo pro segmenty normálního rytmu dosaženo senzitivity 99,17 %, specificity 57,15 %, PPV 87,58 %, přesnosti 88,79 % a F1 skóre 93,01 %, pro segmenty KES se senzitivita rovnala 72,60 %, specificita 99,62 %, PPV 93,58 %, přesnost 97,73 % a F1 skóre 81,76 %. U třídy ostatních arytmií vyšla klasifikace nejhůře, senzitivita dosáhla pouze 40,29 %, specificita 97,35 %, PPV 76,60 %, přesnost 87,25 % a F1 skóre 52,80 %. Výsledky dosažené v této práci jsou detailněji rozebrány v kapitole 5.2 a diskutovány v následující podkapitole.

5.4 Diskuse dosažených výsledků

Tato část je věnována diskusi použité metody, dosažených výsledků, možných limitací a nedostatků navrženého algoritmu a představení možných rozšíření a vylepšení metody.

Prvním podstatným úkolem byl výběr a předzpracování dat a vytvoření trénovací, validační a testovací množiny. Jelikož byly téměř ve všech publikovaných metodách využity záznamy z MIT-BIH AD, pro možnost lepšího srovnání výsledků byly tyto záznamy použity i v této práci. V převážné většině jednodušších metod, které ke klasifikaci nevyužívaly hlubokého učení, a také u některých metod, které ho využívaly, byly EKG signály v rámci přípravy dat filtrovány za účelem snížení šumu. V rámci tvorby této práce byla nejprve snaha signály také filtrovat, ale jelikož bylo při trénování autoenkodéru zjištěno, že vlastností této sítě byla zároveň redukce šumu, bylo od samotné filtrace upuštěno. Filtrace pak nebyla použita ani poté, co bylo učení autoenkodéru vynecháno a celá síť byla rovnou trénována ke klasifikaci. Z výsledků dosažených pro klasifikaci do dvou tříd vyplývá, že síť je k přítomnosti šumu dostatečně robustní a šum tak výsledek klasifikace příliš neovlivní. To ale může být způsobeno i tím, že se od sebe EKG záznam normálního rytmu a KES liší, QRS komplex u KES je rozšířený a má abnormální tvar, a jeho průběh není šumem ovlivněn tolik jako průběh vln P či T. Z výsledků klasifikace do tří tříd lze naopak pozorovat, že naprostá většina úseků síňových extrasystol byla chybně klasifikována jako normální rytmus. Tvar QRS komplexu normálního rytmu a SVES je stejný nebo velmi podobný, pouze vlna P má odlišný tvar. Tento jev ale v signálu ovlivněném šumem prakticky nejde pozorovat. Vynechání filtrace proto mohlo výsledky klasifikace negativně ovlivnit, a jedním z možných vylepšení práce by tak mohla být filtrace šumu v signálech před jejich využitím k trénování sítě.

Dalším krokem bylo rozdělení EKG záznamů na jednotlivé úseky. V naprosté většině metod byly záznamy rozděleny na segmenty stejné délky s kmitem R v přibližně jedné třetině až v jedné polovině každého úseku, kdy celková délka úseku byla nejčastěji 200 až 300 vzorků. Také v této práci měly nejprve všechny segmenty shodnou délku 260 vzorků (při vzorkovací frekvenci 360 Hz odpovídá délce trvání přibližně 0,72 s) s tím, že bylo vybráno vždy 99 vzorků před R kmitem a 160 za ním, ale při bližším zkoumání takto vytvořených záznamů se dalo pozorovat, že kvůli proměnné tepové frekvenci se v některých záznamech objevil i následující R kmit, v jiných naopak byla jen část vlny T nebo P nebo některá vlna chyběla úplně. S ohledem na to byla nakonec délka každého úseku volena podle okamžité tepové frekvence (tzn. podle délky aktuálního RR intervalu) tak, že byl vybrán počet vzorků odpovídající 1/3 délky RR intervalu před R kmitem a počet vzorků odpovídající 2/3 délky RR intervalu za ním. Každý vytvořený úsek měl mít celkovou délku 480 vzorků, pokud tedy byl některý RR interval kratší, byl daný úsek na tuto délku doplněn dvěma různými způsoby. Při prvním způsobu byl úsek doplněn na konci hodnotou posledního vzorku, u druhého způsobu byl úsek doplněn z levé i pravé strany první a poslední hodnotou signálu tak, aby byl R kmit vždy na pozici 160. vzorku.

U druhého způsobu tak byly na rozdíl od prvního všechny R kmity na stejné pozici. Jak vyplynulo z výsledků klasifikace do dvou tříd, první způsob doplnění byl o něco úspěšnější, při klasifikaci do tří tříd proto byl využit pouze tento způsob. Délka úseku 480 vzorků byla vybrána s ohledem na následující zpracování signálu pomocí neuronové sítě, kdy při zvolené architektuře byl každý úsek převeden do vektoru 30 příznaků, které byly poté klasifikovány pomocí LSTM jednotek a plně propojených vrstev. V rámci tvorby práce byly uvažovány i jiné délky, a to 400 vzorků (odpovídá vektoru 25 příznaků), 440 vzorků (vektor 27 příznaků), 460 vzorků (28 příznaků), 500 vzorků (31 příznaků), 520 vzorků (32 příznaků) a 560 vzorků (35 příznaků). Úseky dlouhé 400 vzorků při vzorkovací frekvenci 360 Hz odpovídají úsekům o délce trvání přibližně 1,1 s, a tedy průměrné tepové frekvenci 54 tepů/min, a úseky dlouhé 560 vzorků odpovídají délce 1,56 s a tepové frekvenci přibližně 39 tepů/min. Délka úseku 480 vzorků (odpovídající průměrné tepové frekvenci 45 tepů/min) tak nebyla vybrána z důvodu lepších dosažených výsledků, ale spíše s ohledem na projevy KES v EKG záznamu, aby při segmentaci jednotlivých srdečních úderů nedošlo ke ztrátě důležité informace (např. P nebo T vln). I přesto ale v některých segmentech část vlny P nebo T chyběla, což mohlo mít (zvláště při klasifikaci do tří tříd) negativní vliv na výsledek. Jako další možné vylepšení algoritmu by tak mohla být podrobnější optimalizace a zaměření na tvorbu úseků odpovídajících jednotlivým srdečním úderům.

Po vytvoření úseků byla data rozdělena do trénovací, validační a testovací množiny. Kvůli více realistickým výsledkům a možnosti hodnocení toho, jak síť dokáže zobecňovat při klasifikaci nových záznamů byl k rozdělení dat použit *subject-oriented* přístup. Všechny úseky z jednoho záznamu tak byly přiřazeny pouze do jedné množiny. Jak je vidět z kapitoly 4.2, záznamy se u klasifikace do dvou tříd podařilo do trénovací, validační a testovací množiny rozložit tak, že bylo ve všech množinách přibližně stejné procentuální zastoupení segmentů s oběma typy rytmu. To bylo možné hlavně díky tomu, že 33 záznamů ze 44 obsahovalo alespoň jeden úsek KES, a poměrově podobné rozdělení do množin tak nebyl příliš velký problém. Naopak, jak je zřejmé z kapitoly 4.4, při klasifikaci do tří tříd je rozložení dat jiné v každé množině. Ve validační i testovací množině bylo procentuálně více úseků BLRT, naopak v trénovací množině bylo více SVES i BPRT. Toto nepřesné rozložení je způsobeno hlavně tím, že pouze 4 záznamy ze 44 obsahovaly úderu typu BLRT a 6 záznamů BPRT. Úseků typu SVES bylo obecně k dispozici velmi málo. Takovéto rozložení záznamů, i když velmi rozdílné pro každou množinu, bylo při *subject-oriented* přístupu za daných okolností pravděpodobně nejlepší možné. Do trénovací množiny tak ale byly zařazeny pouze 2 záznamy s úseky typu BLRT a 4 s BPRT, a ve validační i testovací množině pak byl pouze jeden záznam s BLRT a jeden s BPRT. Toto rozložení mohlo být jeden z důvodů, proč síť úspěšně klasifikovala rytmus do dvou tříd, ale do tří už ne, a proč bylo tolik úseků SVES, BLRT i BPRT ve validační i testovací množině označeno jako normální rytmus. Při využití *class-oriented* způsobu by se daly jednotlivé úseky do množin rozdělit mnohem přesněji, ale výsledky

klasifikace by pak byly přehnaně optimistické. Jediné řešení tohoto problému u *subject-oriented* přístupu, a tedy možné vylepšení využití metody, je rozšíření trénovací množiny o co nejvíce dalších záznamů s velkou různorodostí zachycených arytmií a srdečních rytmů.

V rámci práce bylo testováno několik architektur neuronových sítí lišících se kromě rozměrů dat vstupujících do první vrstvy i celkovým počtem a rozložením vrstev a množstvím neuronů v jednotlivých vrstvách. Optimalizaci použité architektury ale bylo věnováno poměrně malé množství času a nebyl na ni kladen příliš velký důraz, v práci použitá architektura sítí byla nakonec velmi podobná architektuře popsané v článku [63]. Dalším možným vylepšením navržené metody je tedy důkladnější zkoušení a zkoumání různých architektur sítí.

Po zvolení vhodné architektury byla celá síť trénována s různým nastavením čtyř hyperparametrů, a to kroku učení, počtu epoch, velikosti dávky dat (*batchsize*) a parametru *rate* v dropout vrstvě, a také s různou normalizací vstupních dat. Tyto hyperparametry pro síť sloužící ke klasifikaci do dvou tříd byly optimalizovány velmi podrobně a pečlivě, byly zkoušeny všechny možné kombinace pro zvolené rozsahy hodnot parametrů. Zde tedy příliš velký prostor pro vylepšení metody pravděpodobně není. U klasifikace do tří tříd však změna hyperparametrů ve stejných rozsazích jako u klasifikace do dvou tříd nevedla k téměř žádnému zlepšení výsledků, a na další podrobnou optimalizaci s jinými rozsahy hodnot už nebyl prostor. Důkladná optimalizace hyperparametrů sítí pro klasifikaci do tří tříd by tedy mohla vést k vylepšení dosažených výsledků. Stejně tak by ke zlepšení výsledků mohla vést optimalizace parametrů, které v rámci práce vůbec měněny nebyly, a to optimalizační funkce a metrika při učení sítě. Místo použité přesnosti by pravděpodobně byla vhodnější metrika F1 skóre.

U klasifikace do dvou tříd bylo dosaženo lepších výsledků při doplnění dat prvním způsobem. U trénovací množiny byla častější chybou falešně pozitivní detekce, z celkového počtu 56769 segmentů bylo celkem 243 úseků (0,43 %) klasifikováno špatně. Naopak u validační i testovací množiny došlo k více falešně negativním detekcím. Z celkového počtu 12386 validačních úseků jich bylo 53 (0,43 %) klasifikováno chybně, stejně tak bylo špatně zařazeno 73 (0,60 %) z 12248 testovacích segmentů. Dá se tak předpokládat, že síť dokáže dobře klasifikovat i data, na kterých se neučila a které vidí poprvé. Výsledky jsou podrobněji shrnuty v kapitole 5.2.1 a tabulkách 5.8 a 5.9. Pro testovací množinu se dosažená senzitivita rovnala 94,87 %, specificita 99,82 %, PPV 98,00 %, přesnost 99,40 % a F1 skóre 96,41 %. Dá se tak říct, že výsledky jsou srovnatelné s výsledky publikovaných metod, i když toto srovnání se nedá považovat za úplně přesné kvůli jinak rozděleným datovým množinám.

U klasifikace do tří tříd došlo k výraznému přeučení sítě, jelikož dobrých výsledků bylo dosaženo pouze pro trénovací množinu dat. Jedno z možných vysvětlení je to, že do třídy ostatní arytmiie byly zařazeny poruchy převodu vzruchu, které se v EKG záznamu projevují velmi odlišně, QRS komplexy i vlny P a T mají různé průběhy a tvary.

S ohledem na tento fakt byla v rámci tvorby práce zkoušena klasifikace do 4 a 5 tříd, kde pro 4 třídy bylo rozdělení na normální rytmus, KES, SVES a raménkové blokády a pro 5 tříd na normální rytmus, KES, SVES, BLRT a BPRT, ale toto rozdělení k lepším výsledkům nevedlo. Jako nejvýraznější vysvětlení se proto jeví nepřesné rozdělení dat do trénovací, validační a testovací množiny a obecný nedostatek trénovacích dat pro ostatní arytmie. Tomuto předpokladu by odpovídal i fakt, že výsledky u testovací množiny byly lepší než pro množinu validační. U validační množiny byly navíc zařazeny špatně téměř všechny úseky BPRT, u testovací naopak většina úseků BLRT. Naprostá většina úseků SVES byla chybně klasifikována jako normální rytmus u obou těchto množin. Další možné příčiny přeučení byly diskutovány v předchozím textu.

Mezi možná rozšíření metody se dá zařadit klasifikace EKG signálů do více tříd, nejlépe do tolika tříd, kolik typů arytmí je v záznamech MIT-BIH AD. Protože však síť nedokázala klasifikovat záznamy ani do tří tříd, před jakýmkoli rozšířením by měla být klasifikace vylepšena. Možné návrhy k optimalizaci metody jsou shrnuty v předchozí diskusi, je to především zaměření na filtraci šumu, tvorbu úseků odpovídajících srdečním úderům, použití většího množství záznamů ke trénování, výběr vhodné architektury sítě a s tím spojené nastavení hyperparametrů. Dalším rozšířením klasifikace srdečního rytmu je pak využití signálů z více svodů.

ZÁVĚR

Cílem práce bylo popsat problematiku komorových extrasystol a dalších patologií a jejich projevů v EKG záznamu, provést literární rešerši metod automatické detekce této arytmie, jednu navrženou metodu realizovat ve vhodném programovém prostředí a otestovat ji na reálných záznamech, tento algoritmus optimalizovat pro dosažení co nejlepších výsledků, a tyto výsledky shrnout, interpretovat a porovnat v kontextu výsledků z jiných studií.

V praktické části práce je realizována metoda klasifikace rytmu srdečních úderů využívající konvolučních neuronových sítí a LSTM jednotek. Jednotlivé srdeční úderý jsou klasifikovány do dvou tříd jako normální rytmus nebo KES i do tří tříd jako normální rytmus, KES a ostatní arytmie. Navržená metoda byla trénována a testována na 44 záznamech z MIT-BIH Arrhythmia Database. Při klasifikaci do dvou tříd bylo na testovacích datech dosaženo senzitivity 94,87 %, specificity 99,82 %, pozitivní prediktivní hodnoty 98,00 %, přesnosti 99,40 % a F1 skóre 96,41 %. Při klasifikaci do tří tříd došlo k výraznému přeučení sítě, u testovací množiny dat se pro segmenty KES senzitivita rovnala 72,60 %, specificita 99,62 %, pozitivní prediktivní hodnota 93,58 %, přesnost 97,73 % a F1 skóre 81,76 %.

Dosažené výsledky jsou v kapitole 5.3 porovnané s výsledky z jiných prací, které byly představeny v kapitole 2. Pro klasifikaci do dvou tříd se výsledky této práce dají považovat za srovnatelné s jinými metodami využívajícími hluboké učení a lepší než výsledky dosažené u jednodušších metod, toto srovnání se však dá považovat pouze za orientační kvůli jinak rozděleným datovým množinám a jiným přístupům k rozdělení dat.

Z výsledků této práce vyplývá, že využití hlubokých neuronových sítí, přesněji spojení konvoluční neuronové sítě a LSTM jednotky, má pro detekci KES (a dalších arytmií) v EKG záznamech velký potenciál. Jak ale bylo vidět u klasifikace do tří tříd, pro úspěšné natrénování sítě je naprosto klíčové dostatečné množství náležitě předzpracovaných a připravených EKG záznamů, výběr vhodné architektury sítě i nastavení hyperparametrů.

LITERATURA

- [1] TROJAN, Stanislav. *Lékařská fyziologie*. Vyd. 4., přeprac. a dopl. Praha: Grada, 2003. ISBN 80-247-0512-5.
- [2] LANGMEIER, Miloš. *Základy lékařské fyziologie*. 1. vydání. Praha: Grada, 2009. ISBN 978-80-247-2526-0.
- [3] FONTANA, Josef, Josef KROUPA, Patrik MAĎA, Peter IVÁK a kolektiv. *Srdce. Funkce buněk a lidského těla: Multimediální skripta* [online]. Praha: 3. lékařská fakulta Univerzity Karlovy, 2016 [cit. 2020-12-27]. Dostupné z: <http://fbt.cz/skripta/x-srdce-a-obeh-krve/1-srdce/>
- [4] SILBERNAGL, Stefan a Agamemnon DESPOPOULOS. *Atlas fyziologie člověka: překlad 8. německého vydání*. 4. české vydání. Praha: Grada Publishing, 2016. ISBN 978-80-247-4271-7.
- [5] SEDMERA, David a František VOSTÁREK. Moderní pohled na převodní systém srdeční. *Časopis lékařů českých* [online]. Praha: Česká lékařská společnost J. E. Purkyně, 2017, **156**(8), 417-421 [cit. 2020-12-27]. ISSN 1805-4420. Dostupné z: <https://www.prolekare.cz/casopisy/casopis-lekaru-ceskych/2017-8/moderni-pohled-na-prevodni-system-srdecni-62750>
- [6] HAMAN, Petr. Srdeční převodní soustava. *Výukový web EKG* [online]. Plzeň [cit. 2020-12-27]. Dostupné z: <http://ekg.kvalitne.cz/system.htm>
- [7] GUYTON, Arthur a John HALL. *Textbook of medical physiology*. 11th ed. Philadelphia: Elsevier Saunders, c2006. ISBN 0-7216-0240-1.
- [8] HAMPTON, John R. EKG stručně, jasně, přehledně. Překlad 7. vydání. Praha: Grada, 2013. ISBN 978-80-247-4246-5.
- [9] KNECHTEL, Maureen A. *EKGs for the Nurse Practitioner and Physician Assistant*. Second Edition. New York: Springer Publishing Company, 2017. ISBN 978-0-8261-6888-7.
- [10] HABERL, Ralph. *EKG do kapsy*. Překlad 4. vydání. Praha: Grada, 2012. ISBN 978-80-247-4192-5.
- [11] COVIELLO, Jessica Shank, ed. *ECG Interpretation Made Incredibly Easy*. 5th Edition. Philadelphia, USA: Wolters Kluwer/Lippincott Williams & Wilkins Health, 2011. ISBN 978-1-60831-289-4.
- [12] THALER, Malcolm S. *The Only EKG Book You'll Ever Need*. 9th Ed. Philadelphia: Wolters Kluwer/Lippincott Williams & Wilkins Health, 2018. ISBN 978-1-49-637723-4.
- [13] BENNETT, David. *Srdeční arytmie: praktické poznámky k interpretaci a léčbě*. Překlad 8. vydání. Praha: Grada, 2014. ISBN 978-80-247-5134-4.
- [14] SIMPSON, Ross J., Wayne E. CASCIO, Pamela J. SCHREINER, Richard S. CROW, Pentti M. RAUTAHARJU a Gerardo HEISS. Prevalence of premature ventricular contractions in a population of African American and white men and women: The Atherosclerosis Risk in Communities (ARIC) study. *American Heart*

- Journal* [online]. 2002, **143**(3), 535-540 [cit. 2020-12-29]. ISSN 00028703. Dostupné z: doi:10.1067/mhj.2002.120298
- [15] MORAVEC, Ondřej, Tomáš SKÁLA a Petr HEINC. Kdy a jak léčit komorové extrasystoly v interní ambulanci. *Interní medicína pro praxi* [online]. 2014, **16**(3), 123-126 [cit. 2020-12-29]. Dostupné z: <https://www.internimedica.cz/pdfs/int/2014/03/10.pdf>
- [16] HEINC, Petr. Kdy léčit komorové extrasystoly. *Interní medicína pro praxi* [online]. 2007, **9**(10), 448-452 [cit. 2020-12-29]. Dostupné z: <https://www.internimedica.cz/pdfs/int/2007/10/07.pdf>
- [17] BURNS, Ed. Premature Ventricular Complex (PVC). In: *Life in the Fast Lane: ECG Library* [online]. Austrálie, 2020 [cit. 2020-12-29]. Dostupné z: <https://litfl.com/premature-ventricular-complex-pvc-ecg-library/>
- [18] BYTEŠNÍK, Jan, Petr PAŘÍZEK, Dan WICHTERLE, Josef KAUTZNER a Petr NEUŽIL. Komorové arytmie. *Cor et Vasa* [online]. 2011, **53**(1), 53-77 [cit. 2020-12-30]. ISSN 00108650. Dostupné z: doi:10.33678/cor.2011.194
- [19] BLAHÚT, Peter. Komorová extrasystola. In: *TECHMED: EKG & Arytmológia* [online]. Bratislava, 2017 [cit. 2020-12-30]. Dostupné z: <https://www.techmed.sk/komorova-extrasystola/>
- [20] GEITER, Henry B. *E-Z ECG Rhythm Interpretation*. Philadelphia, USA: F. A. Davis Company, 2007. ISBN 978-0-8036-1043-9.
- [21] ITTATIRUT, Supat, Apiwat LEK-UTHAI a Arporn TEERAMONGKONRASMEE. Detection of Premature Ventricular Contraction for real-time applications. *2013 10th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology* [online]. Krabi: IEEE, 2013, **10**, 1-5 [cit. 2020-11-07]. ISBN 978-1-4799-0545-4. Dostupné z: doi:10.1109/ECTICon.2013.6559531
- [22] CAI, Zhipeng, Jianqing LI, Alistair E W JOHNSON, Xiangyu ZHANG, Qin SHEN, Jian ZHANG a Chengyu LIU. Rule-based rough-refined two-step-procedure for real-time premature beat detection in single-lead ECG. *Physiological Measurement* [online]. 2020, **41**(5), 1-22 [cit. 2020-11-09]. ISSN 1361-6579. Dostupné z: doi:10.1088/1361-6579/ab87b4
- [23] GHORBANI AFKHAMI, Rashid, Ghanbar AZARNIA a Mohammad Ali TINATI. Cardiac arrhythmia classification using statistical and mixture modeling features of ECG signals. *Pattern Recognition Letters* [online]. 2016, **70**, 45-51 [cit. 2020-11-07]. ISSN 01678655. Dostupné z: doi:10.1016/j.patrec.2015.11.018
- [24] YE, Can, B. V. K. Vijaya KUMAR a M. T. COIMBRA. Heartbeat Classification Using Morphological and Dynamic Features of ECG Signals. *IEEE Transactions on Biomedical Engineering* [online]. 2012, **59**(10), 2930-2941 [cit. 2020-11-23]. ISSN 0018-9294. Dostupné z: doi:10.1109/TBME.2012.2213253
- [25] MARTIS, Roshan Joy, U. Rajendra ACHARYA, K.M. MANDANA, A.K. RAY a Chandan CHAKRABORTY. Application of principal component analysis to ECG signals for automated diagnosis of cardiac health. *Expert Systems with Applications*

- [online]. 2012, **39**(14), 11792-11800 [cit. 2020-11-07]. ISSN 09574174. Dostupné z: doi:10.1016/j.eswa.2012.04.072
- [26] ZAREI, Roozbeh, Jing HE, Guangyan HUANG a Yanchun ZHANG. Effective and efficient detection of premature ventricular contractions based on variation of principal directions. *Digital Signal Processing* [online]. 2016, **50**, 93-102 [cit. 2020-11-07]. ISSN 10512004. Dostupné z: doi:10.1016/j.dsp.2015.12.002
- [27] MALEK, Abdel Salam, Ashraf ELNAHRAWY, Hamed ANWAR a Mohamed NAEEM. Automated detection of premature ventricular contraction in ECG signals using enhanced template matching algorithm. *Biomedical Physics & Engineering Express* [online]. 2020, **6**(1), 1-12 [cit. 2020-11-08]. ISSN 2057-1976. Dostupné z: doi:10.1088/2057-1976/ab6995
- [28] LIU, Chengyu, Peng LI, Yatao ZHANG, Yuan ZHANG, Changchun LIU a Shoushui WEI. A Construction Method of Personalized ECG Template and Its Application in Premature Ventricular Contraction Recognition for ECG Mobile Phones. In: *World Congress on Medical Physics and Biomedical Engineering May 26-31, 2012, Beijing, China* [online]. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, s. 585-588 [cit. 2020-11-08]. IFMBE Proceedings. ISBN 978-3-642-29304-7. Dostupné z: doi:10.1007/978-3-642-29305-4_153
- [29] JUN, Tae Joon, Hyun Ji PARK, Nguyen Hoang MINH, Daeyoung KIM a Young-Hak KIM. Premature Ventricular Contraction Beat Detection with Deep Neural Networks. In: *2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)* [online]. Anaheim, CA: IEEE, 2016, s. 859-864 [cit. 2020-11-08]. ISBN 978-1-5090-6167-9. Dostupné z: doi:10.1109/ICMLA.2016.0154
- [30] SANNINO, G. a G. DE PIETRO. A deep learning approach for ECG-based heartbeat classification for arrhythmia detection. *Future Generation Computer Systems* [online]. 2018, **86**(9), 446-455 [cit. 2020-11-08]. ISSN 0167739X. Dostupné z: doi:10.1016/j.future.2018.03.057
- [31] KHALAF, Aya F., Mohamed I. OWIS a Inas A. YASSINE. A novel technique for cardiac arrhythmia classification using spectral correlation and support vector machines. *Expert Systems with Applications* [online]. 2015, **42**(21), 8361-8368 [cit. 2020-11-09]. ISSN 09574174. Dostupné z: doi:10.1016/j.eswa.2015.06.046
- [32] ATANASOSKI, V., M. D. IVANOVIC, M. MARINKOVIC, G. GLIGORIC, B. BOJOVIC, A. V. SHVILKIN a J. PETROVIC. Unsupervised Classification of Premature Ventricular Contractions Based on RR Interval and Heartbeat Morphology. In: *2018 14th Symposium on Neural Networks and Applications (NEUREL)* [online]. Bělehrad: IEEE, 2018, s. 1-6 [cit. 2020-11-22]. ISBN 978-1-5386-6974-7. Dostupné z: doi:10.1109/NEUREL.2018.8586997
- [33] GORDON, Max a Cranos WILLIAMS. PVC Detection Using a Convolutional Autoencoder and Random Forest Classifier. *Pacific Symposium on Biocomputing*

- [online]. World Scientific, 2019, **24**, 42-53 [cit. 2020-11-21]. ISBN 978-981-327-981-0. Dostupné z: doi:10.1142/9789813279827_0005
- [34] OCHIAI, Keiichi, Shu TAKAHASHI a Yusuke FUKAZAWA. Arrhythmia Detection from 2-lead ECG using Convolutional Denoising Autoencoders. *KDD'18 Deep Learning Day* [online]. London, UK, 2018, **24** [cit. 2020-11-30]. Dostupné z: https://www.kdd.org/kdd2018/files/deep-learning-day/DLDay18_paper_16.pdf
- [35] YANG, Jianli, Yang BAI, Feng LIN, Ming LIU, Zengguang HOU a Xiuling LIU. A novel electrocardiogram arrhythmia classification method based on stacked sparse auto-encoders and softmax regression. *International Journal of Machine Learning and Cybernetics* [online]. 2018, **9**(10), 1733-1740 [cit. 2020-12-01]. ISSN 1868-8071. Dostupné z: doi:10.1007/s13042-017-0677-5
- [36] ACHARYA, U. Rajendra, Shu Lih OH, Yuki HAGIWARA, Jen Hong TAN, Muhammad ADAM, Arkadiusz GERTYCH a Ru San TAN. A deep convolutional neural network model to classify heartbeats. *Computers in Biology and Medicine* [online]. 2017, **89**, 389-396 [cit. 2020-11-23]. ISSN 00104825. Dostupné z: doi:10.1016/j.compbimed.2017.08.022
- [37] ZHANG, Yuwei, Jianxun YU, Yuan ZHANG, Chengyu LIU a Jianqing LI. A Convolutional Neural Network for Identifying Premature Ventricular Contraction Beat and Right Bundle Branch Block Beat. In: *2018 International Conference on Sensor Networks and Signal Processing (SNSP)* [online]. Si-an, Čína: IEEE, 2018, s. 158-162 [cit. 2020-11-23]. ISBN 978-1-5386-7413-0. Dostupné z: doi:10.1109/SNSP.2018.00037
- [38] LI, Dan, Jianxin ZHANG, Qiang ZHANG a Xiaopeng WEI. Classification of ECG signals based on 1D convolution neural network. In: *2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom)* [online]. Ta-lien: IEEE, 2017, s. 1-6 [cit. 2020-11-23]. ISBN 978-1-5090-6704-6. Dostupné z: doi:10.1109/HealthCom.2017.8210784
- [39] STEWART, Matthew. Simple Introduction to Convolutional Neural Networks. *Towards Data Science* [online]. February 27, 2019 [cit. 2021-01-04]. Dostupné z: <https://towardsdatascience.com/simple-introduction-to-convolutional-neural-networks-cdf8d3077bac>
- [40] BROWNLEE, Jason. How Do Convolutional Layers Work in Deep Learning Neural Networks? *Machine Learning Mastery* [online]. April 17, 2019 [cit. 2021-01-04]. Dostupné z: <https://machinelearningmastery.com/convolutional-layers-for-deep-learning-neural-networks/>
- [41] YAMASHITA, Rikiya, Mizuho NISHIO, Richard Kinh Gian DO a Kaori TOGASHI. Convolutional neural networks: an overview and application in radiology. *Insights into Imaging* [online]. 2018, **9**(4), 611-629 [cit. 2021-01-05]. ISSN 1869-4101. Dostupné z: doi:10.1007/s13244-018-0639-9
- [42] SAHA, Sumit. A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way. *Towards Data Science* [online]. December 15, 2018 [cit. 2021-01-04].

- Dostupné z: <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>
- [43] PRAHBU, Raghav. Understanding of Convolutional Neural Network (CNN) — Deep Learning. *Medium* [online]. March 4, 2018 [cit. 2021-01-04]. Dostupné z: <https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148>
- [44] AMIDI, Afshine a Shervine AMIDI. Convolutional Neural Networks cheatsheet. *Stanford University* [online]. November 26, 2018 [cit. 2021-01-05]. Dostupné z: <https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-convolutional-neural-networks>
- [45] SHARMA, Avinash. Understanding Activation Functions in Neural Networks. *Medium: The Theory of Everything* [online]. March 30, 2017 [cit. 2021-01-05]. Dostupné z: <https://medium.com/the-theory-of-everything/understanding-activation-functions-in-neural-networks-9491262884e0>
- [46] SHARMA, Sagar. Activation Functions in Neural Networks. *Towards Data Science* [online]. September 6, 2017 [cit. 2021-01-05]. Dostupné z: <https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6>
- [47] MALADKAR, Kishan. Types Of Activation Functions In Neural Networks And Rationale Behind It. *Analytics India Magazine* [online]. 2018 [cit. 2021-01-05]. Dostupné z: <https://analyticsindiamag.com/most-common-activation-functions-in-neural-networks-and-rationale-behind-it/>
- [48] BHATTACHARYYA, Jayita. Activation Functions in Neural Networks: An Overview. In: *Analytics India Magazine* [online]. November 6, 2020 [cit. 2021-01-05]. Dostupné z: <https://analyticsindiamag.com/activation-functions-in-neural-network/>
- [49] SHUKLA, Rishabh. L1 vs. L2 Loss function. *Github* [online]. July 28, 2015 [cit. 2021-04-16]. Dostupné z: <http://rishy.github.io/ml/2015/07/28/l1-vs-l2-loss/>
- [50] DWIVEDI, Rohit. What Are Different Loss Functions Used as Optimizers in Neural Networks? *Analytic Steps* [online]. June 17, 2020 [cit. 2021-04-16]. Dostupné z: <https://www.analyticssteps.com/blogs/what-are-different-loss-functions-used-optimizers-neural-networks>
- [51] VERMA, Shiva. Understanding different Loss Functions for Neural Networks. *Towards Data Science* [online]. June 20, 2019 [cit. 2021-04-16]. Dostupné z: <https://towardsdatascience.com/understanding-different-loss-functions-for-neural-networks-dd1ed0274718>
- [52] KOECH, Kiprono Elijah. Cross-Entropy Loss Function. *Towards Data Science* [online]. October 2, 2020 [cit. 2021-04-16]. Dostupné z: <https://towardsdatascience.com/cross-entropy-loss-function-f38c4ec8643e>
- [53] DOSHI, Sanket. Various Optimization Algorithms For Training Neural Network. *Towards Data Science* [online]. January 13, 2019 [cit. 2021-5-11]. Dostupné z:

- <https://towardsdatascience.com/optimizers-for-training-neural-network-59450d71caf6>
- [54] LUHANIWAL, Vikash Raj. Why Gradient descent isn't enough: A comprehensive introduction to optimization algorithms in neural networks. *Towards Data Science* [online]. May 7, 2019 [cit. 2021-5-11]. Dostupné z: <https://towardsdatascience.com/why-gradient-descent-isnt-enough-a-comprehensive-introduction-to-optimization-algorithms-in-59670fd5c096>
- [55] RUDER, Sebastian. An overview of gradient descent optimization algorithms. *Sebastian Ruder - ruder.io* [online]. January 19, 2016 [cit. 2021-5-11]. Dostupné z: <https://ruder.io/optimizing-gradient-descent/index.html>
- [56] HOCHREITER, Sepp a Jürgen SCHMIDHUBER. Long Short-Term Memory. *Neural Computation* [online]. 1997, **9**(8), 1735-1780 [cit. 2021-01-04]. ISSN 0899-7667. Dostupné z: doi:10.1162/neco.1997.9.8.1735
- [57] OLAH, Christopher. Understanding LSTM Networks. *Colah's blog* [online]. August 27, 2015 [cit. 2021-01-04]. Dostupné z: <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>
- [58] PHI, Michael. Illustrated Guide to LSTM's and GRU's: A step by step explanation. *Towards Data Science* [online]. September 24, 2018 [cit. 2021-01-04]. Dostupné z: <https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation-44e9eb85bf21>
- [59] KANG, Eugene. Long Short-Term Memory (LSTM): Concept. *Medium* [online]. September 2, 2017 [cit. 2021-01-04]. Dostupné z: <https://medium.com/@kangeugine/long-short-term-memory-lstm-concept-cb3283934359>
- [60] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation* **101**(23):e215-e220 [Circulation Electronic Pages; <http://circ.ahajournals.org/content/101/23/e215.full>]; 2000 (June 13).
- [61] MOODY, George B. a Roger G. MARK. The impact of the MIT-BIH Arrhythmia Database. *IEEE Engineering in Medicine and Biology Magazine* [online]. 2001, **20**(3), 45-50 [cit. 2021-01-02]. ISSN 07395175. Dostupné z: doi:10.1109/51.932724
- [62] MOODY, George B. MIT-BIH Arrhythmia Database Directory: Introduction. *PhysioNet* [online]. Boston: Harvard-MIT Division of Health Sciences and Technology, Biomedical Engineering Center, 1997 [cit. 2021-01-02]. Dostupné z: <https://archive.physionet.org/physiobank/database/html/mitdbdir/intro.htm>
- [63] YILDIRIM, Ozal, Ulas Baran BALOGLU, Ru-San TAN, Edward J. CIACCIO a U. Rajendra ACHARYA. A new approach for arrhythmia classification using deep coded features and LSTM networks. *Computer Methods and Programs in*

Biomedicine [online]. 2019, **176**, 121-133 [cit. 2021-01-02]. ISSN 01692607.
Dostupné z: doi:10.1016/j.cmpb.2019.05.004

SEZNAM SYMBOLŮ A ZKRATEK

Acc	přesnost
AD	Arrhythmia Database
AV	atrioventrikulární
BCE	binární křížová entropie
BIH	Beth Israel Hospital
BLRT	blokáda levého raménka Tawarova
BPRT	blokáda pravého raménka Tawarova
CAE	konvoluční autoenkodér
CCE	kategorická křížová entropie
CNN	konvoluční neuronová síť
DNN	hluboká neuronová síť
DWT	diskrétní vlnková transformace
EKG	elektrokardiogram
EM	expectation-maximization
FN	falešně negativní
FP	falešně pozitivní
GMM	Gaussian mixture model
ICA	analýza nezávislých komponent
ICHS	ischemická choroba srdeční
KES	komorová extrasystola
LSTM	long-short term memory
MIT	Massachusetts Institute of Technology
MSE	střední kvadratická odchylka
NN	neuronová síť
PCA	analýza hlavních komponent
PPV	pozitivní prediktivní hodnota
RBF	radiální bázová funkce
SA	sinoatriální
Se	senzitivita
Sp	specifická
SVES	supraventrikulární extrasystola
SVM	metoda podpůrných vektorů
TN	pravdivě negativní
TP	pravdivě pozitivní
VAT	doba aktivace komor