

# Klasifikacija zapisa fonokardiograma na normalne i abnormalne

Josip Kiralj

Prirodoslovno matematički fakultet Zagreb, jokiral@student.math.hr

Dajana Jerončić

Prirodoslovno matematički fakultet Zagreb, dajeron@student.math.hr

**Abstract** – Fonokardiogram (PCG) se koristi kao dijagnostički test za evaluiranje hemodinamičkog stanja te za detektiranje kardiovaskularnih bolesti. Cilj 2016 PhysionetNet/Cinc Challenge-a je upravo klasifikacija PCG zapisa otkucaja srca na normalne i abnormalne.

Problem smo prvo pristupili jednostavnijom metodom – logističkom regresijom koju smo proveli za 20 vremenskih značajki dobivenih Springerovom segmentacijom. Dobiveni rezultat je na testnom setu imao specifičnost 67,24%, osjetljivost 76,41%, što daje konačnu mjeru uspješnosti 71,83%. Drugi pristup se bazira na dubokim konvolucijskim neuronskim mrežama i koeficijentima dobivenih MFCC analizom koji je polučio vidno bolje rezultate – specifičnost je iznosila 86,7%, osjetljivost 77,8%, što je na kraju mjera uspješnosti u iznosu od 82,25%.

**Index Terms** – Heartbeat classification, Machine learning, Phonocardiogram

## 1. Uvod

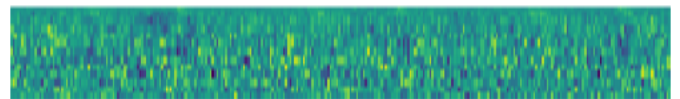
Klasifikacija zvučnih zapisa otkucaja srca odnosi se na problem binarne klasifikacije snimaka otkucaja na normalne i abnormalne. Prema podacima Svjetske zdravstvene organizacije (WHO), kardiovaskularne bolesti primarni su uzrok smrti – više ljudi umre od njih nego od bilo kojeg drugog uzroka. Procijenjeno je da je 2004. 17,1 milijun ljudi umrlo od kardiovaskularnih bolesti, što predstavlja 29% svih umrlih, od čega je za 7,2 milijuna njih uzrok smrti bila neka od koronarnih bolesti srca. Zbog toga je potrebno naći metodu kojom bi se pravovremeno uspjele dijagnosticirati anomalije kod otkucaja srca.

U ovom radu pokušali smo riješiti problem predstavljen na web stranici physionet[1]. Na navedenom natjecanju najveću uspješnost postigla je metoda koja je koristila ansambl AdaBoost-a i konvolucijskih neuronskih mreža[2], te je uspješnost tog rješenja iznosila 86,02%, što je srednja vrijednost specifičnosti i osjetljivosti.

Naš prvi pokušaj se zasniva na jednostavnoj metodi klasifikacije – logističkoj regresiji. Iako jednostavna, metoda je dala rezultate od 78,83% i to samo koristeći vremenske značajke. Druga metoda se zasniva na dubokim konvolucijskim neuronskim mrežama i MFCC koeficijentima te je polučila rezultat od 82,25%.

### 1.1. Prvotna ideja

Prvotna ideja je bila iskoristiti heatmap dobiven MFCC analizom[2] za treniranje konvolucijske neuronske mreže za prepoznavanje slika.

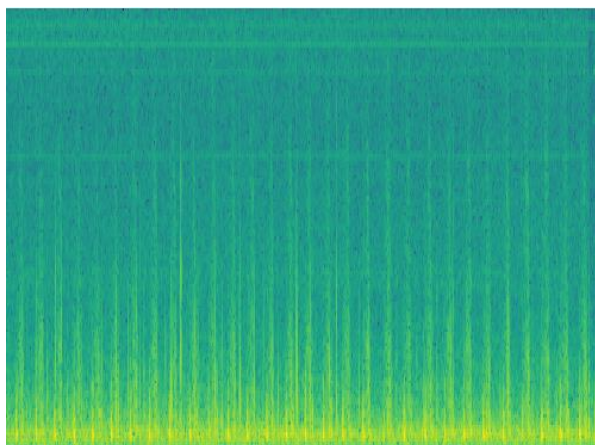


Slika 1. Primjer heatmap-a dobivenog MFCC analizom

U tu svrhu, korištenjem segmentacije smo odredili početak prvog otkucaja srca (S1) te zatim pomoću MFCC analize za 300 prozora trajanja 25ms i razmaka 10ms stvorili heatmap. Nakon toga koristeći konvolucijsku neuronsku mrežu smo pokušali klasificirati dobivene slike, međutim neuspješno, s obzirom da je za razne parametre, balansiranošću skupa i optimizatore, model klasificirao sve zapise u istu klasu. Na slici 1. se nalazi primjer takvog heatmap-a.

Također smo pokušali i za spektrograme (Slika 2.), međutim ni ti pokušaji nisu polučili uspjeha. Zbog toga smo pribjegli jednostavnoj metodi – logističkoj regresiji, nakon koje smo ipak za bolje rezultate koristili metodu zasnovanu na dubokim konvolucijskim neuronskim mrežama.

Lipanj 30, 2018, Zagreb

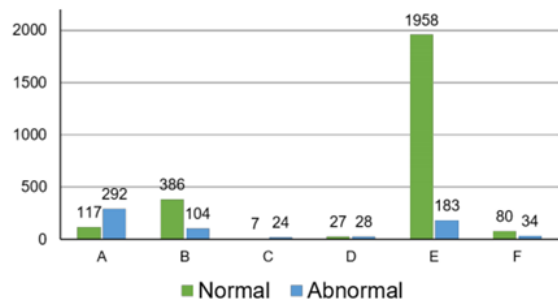


Slika 2. Primjer spektrograma

## 2. Skup podataka

Snimke otkucaja srca prikupljene su kako iz kliničkih, tako i iz nekliničkih okruženja, a sastoje se od snimaka zdravih pojedinaca te bolesnih pacijenata. Skup podataka podijeljen je na šest dijelova (A do F) s ukupno 3 240 snimaka koje traju od 5 do nešto malo više od 120 sekundi Slika 3. prikazuje histogram na kojem se vidi broj snimaka po pojedinom dijelu A do F.

Snimke su snimane na različitim dijelovima tijela (njih 9), od kojih su najčešći: područje aorte, pluća, trikuspidalnog zaliska te mitralnog zaliska. Za bolesne pacijente nije specificirano od čega točno boluju, iako to može biti cijeli niz bolesti, ali tipično to su defekti srčanog ventila te bolesti koronarnih arterija. Od svakog pacijenta može biti najviše 6 snimaka. Frekvencije uzorkovanja su 2000Hz te su snimke dane u .wav formatu, gdje svaka sadrži samo jedan „PCG lead“ (zapis fonokardiograma).



Slika 3. Histogram normalnih i abnormalnih zapisa po dijelovima A do F

Iako se skup podataka izvorno sastojao od 3 240 snimaka, tijekom natjecanja natjecatelji su dobili uputu da određene snimke pomaknu iz skupa podataka s obzirom da ne sadržavaju snimke otkucaja srca ili se pak u njima više čuju zvukovi ostalih organa u tijelu nego samih otkucaja.

Osim toga, mnoge snimke su označene kao “noisy”, što znači da je njihova klasifikacije otežana. Upravo zbog toga je ovaj physionet-ov natječaj od velikog značaja – omogućava klasifikaciju otkucaja srca za koje nije potrebna snimka u kontroliranim uvjetima kao kod elektrokardiograma.

Također je iz histograma primjetno kako je broj normalnih i abnormalnih zapisa otkucaja srca neujednačen, otprilike u omjeru 4:1. Zbog toga se pojavljuje problem nebalansiranosti skupa što uvelike otežava klasifikaciju.

Jedan od problema s kojim smo se susreli je taj što je broj snimaka abnormalnih otkucaja tek nešto veći od 660, dok je normalnih preko 2570. Ako bi uspjeli donekle izbalansirati training set, premalo abnormalnih zapisa bi mogli koristiti u validation te test setu. Međutim ako smo željeli da u test setu broj zapisa za provjeru bude otprilike podjednak, tada bismo imali premalo abnormalnih zapisa za treniranje i validaciju.

Zbog toga smo prvo proveli undersampling. Odvojili smo sve snimke normalnih otkucaja srca iz svih dijelova skupa podataka A-F od onih abnormalnih. Odlučili smo ostaviti sve abnormalne zapise, njih 665, te nasumičnim odabirom odabrati 994 normalna zapisa, nakon čega jer omjer bio ipak prihvatljivijih 3:2 u korist normalnih. Za kraj smo odvojili test set koji smo koristili pri testiranju podataka – taj test set se sastojao od 131 zapisa abnormalnih otkucaja te 136 zapisa normalnih otkucaja, ili približno u omjeru 1:1.

Nakon definiranog training i test seta, za validacijski set smo se odlučili na 20% training seta.

Valja napomenuti kako testni set natjecanja nije objavljen. Prijavljeni natjecatelji imali su na raspolaganju sve podatke za treniranje, te je čak navedeno u uputama da je 300 zapisa danog skupa podataka jedan dio (10%) pravog, većeg dataset-a na kojem će se evaluirati rješenja. Dakle natjecatelji su trenirali algoritme na jednom dijelu skupa podataka na kojem su se i testirala njihova rješenja.

## 3. Mjera uspješnosti

Mjera uspješnosti koju smo koristili u radovima je aritmetička sredina osjetljivosti (eng. sensitivity, recall) te specifičnosti (eng. specificity).

		Predicted	
		Negative	Positive
Actual	Negative	TN	FP
	Positive	FN	TP

Tablica 1. Matrica konfuzije

U našem konkretnom problemu, pozitivnim smatramo klasifikaciju da je snimka otkucaja abnormalna. Shodno tome, formule pomoću kojih računamo uspješnost algoritama su sljedeće:

$$Se = \frac{TP}{TP + FN} \quad (1)$$

$$Sp = \frac{TN}{TN + FP} \quad (2)$$

Te konačna mjera uspješnosti je dana izrazom:

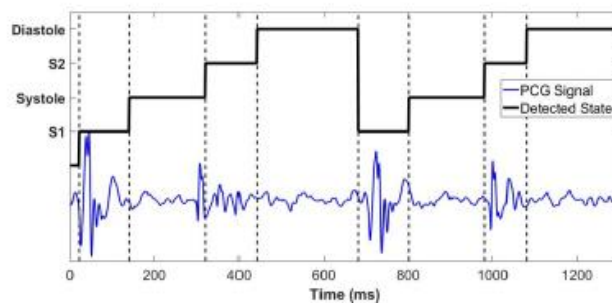
$$Macc = \frac{Se + Sp}{2} \quad (3)$$

## 4. Klasifikacija logističkom regresijom

Za provođenje logističke regresije koristili smo vremenske značajke dobivene nakon segmentacije zvučnih zapisa. Korištena je Springerova metoda[3] za segmentaciju, pomoću koje su određeni počeci srčanih tonova – S1, sistole, S2 i dijastole.

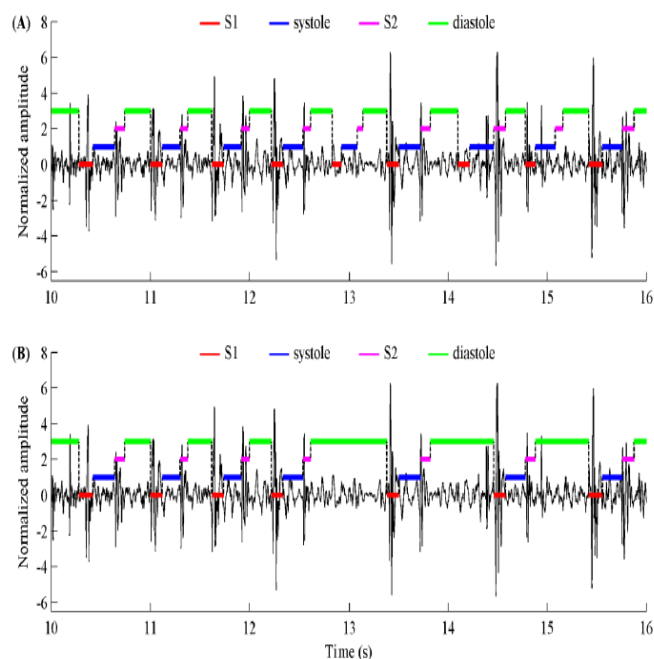
### 4.1. Segmentacija

Springerova metoda[3] zasniva se na skrivenom semi-Markovljevom modelu koji je proširen korištenjem logističke regresije za procjenu vjerojatnosti emisije. Na taj način se pokušao riješiti problem točne segmentacije šumovitih snimaka prikupljenih u svakodnevnim okruženjima.



Slika 4. Srčani tonovi dobiveni Springerovom segmentacijom

Slika 5. prikazuje četiri moguća stanja označena pomoću Springerovog algoritma. Ipak, zbog velikog broja šumovitih snimki, automatsko određivanje stanja pomoću Springerovog algoritma je rezultiralo velikim brojem snimaka koji su bili pogrešno označeni. Zbog toga je tijekom natjecanja objavljen skup .mat datoteka koje su sadržavale ručno popravljene anotacije. Upravo te anotacije smo i koristili pri određivanju značajki. Na slici 3. se može vidjeti razlika u segmentaciji dobivenoj automatski Springerovom metodom (A) te onoj dobivenoj nakon ručno popavljenih anotacija.



Slika 5. Četiri stanja otkucaja srca za isti zapis

(A) Anotacije dobivene automatski (Springerovim algoritmom)

(B) Stanja nakon ručno popavljenih anotacija

## 4.2. Određivanje značajki

Pomoću dobivenih početaka stanja S1, sistole, S2 i dijastole izračunali smo 20 vremenskih značajki. Prije njihovog navođenja napomenimo da je RR interval vremenski interval između dva uzastopna otkucaja srca. Značajke su sljedeće:

- m\_RR – srednja vrijednost RR intervala
- sd\_RR – standardna devijacija RR intervala
- m\_IntS1 – srednja vrijednost S1 intervala
- sd\_IntS1 – standardna devijacija S1 intervala
- m\_IntS2 – srednja vrijednost S2 intervala
- sd\_IntS2 – standardna devijacija S2 intervala
- m\_IntSys – srednja vrijednost sistoličkih intervala
- sd\_IntSys – standardna devijacija sistoličkih intervala
- m\_IntDia – srednja vrijednost dijastoličkih intervala
- sd\_IntDia – standardna devijacija dijastoličkih intervala
- m\_Ratio\_SysRR – srednja vrijednost omjera sistoličkih i RR intervala za svaki otkucaj
- sd\_Ratio\_SysRR – standardna devijacija omjera sistoličkih i RR intervala
- m\_Ratio\_DiaRR – srednja vrijednost omjera dijastoličkih i RR intervala
- sd\_Ratio\_DiaRR – standardna devijacija omjera dijastoličkih i RR intervala
- m\_Ratio\_SysDia – srednja vrijednost omjera sistoličkih i dijastoličkih intervala
- sd\_Ratio\_SysDia – standardna devijacija omjera sistoličkih i dijastoličkih intervala
- m\_Amp\_SysS1 – srednja vrijednost omjera srednje vrijednosti absolutne amplitude tijekom sistole u odnosu na onu tijekom S1
- sd\_Amp\_SysS1 – standardna devijacija omjera srednje vrijednosti absolutne amplitude tijekom sistole u odnosu na onu tijekom S1
- m\_Amp\_DiaS2 – srednja vrijednost omjera srednje vrijednosti absolutne amplitude tijekom dijastole u odnosu na onu tijekom S2
- sd\_Amp\_DiaS2 – standardna devijacija omjera srednje vrijednosti absolutne amplitude tijekom dijastole u odnosu na onu tijekom S2

## 4.3. Logistička regresija

Nakon određivanja 20 značajki za svaki zapis, grupirali smo zapise u dvije grupe: normalne i abnormalne. Računanjem srednjih vrijednosti značajki za svaku grupu vidljivo je da se razlikuju za neke značajke, pogotovo je

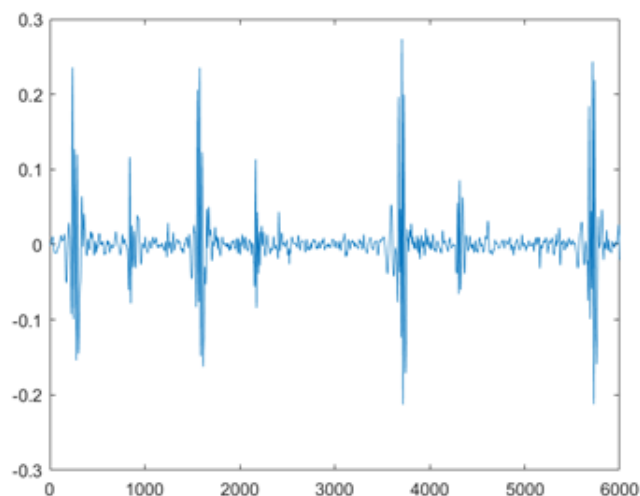
primjetno za značajke koje se odnose na amplitudu. Na slici 4. prvi redak se odnosi na grupu normalnih zapisa, dok drugi redak na grupu abnormalnih zapisa.

sd\_Ratio\_SysDia m\_Amp\_SysS1 sd\_Amp\_SysS1 m\_Amp\_DiaS2 sd\_Amp\_DiaS2

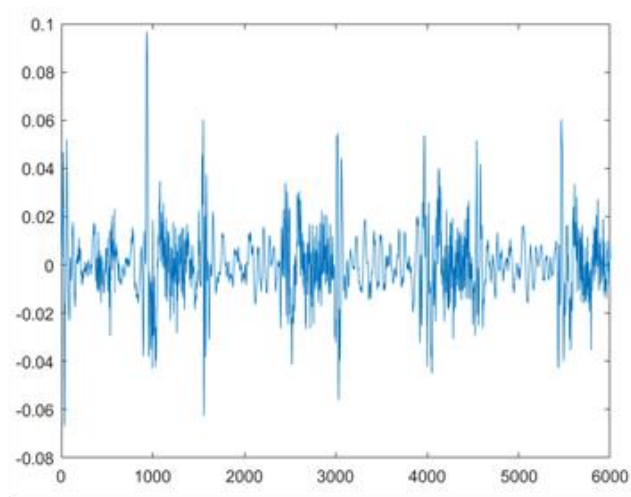
0.045670	1.181521	4.268058	3.205700	0.869203
0.054902	2.004726	7.977303	2.292191	0.781749

Slika 6. Srednja vrijednost značajki za grupirane zapise

Razlika u amplitudama se može vidjeti i u sljedećim prikazima zvučnih valova (Slike 7. i 8.) za normalni te abnormalni otkucaj srca.



Slika 7. Prikaz zvučnog vala kod normalnog otkucaja



Slika 8. Prikaz zvučnog vala kod abnormalnog otkucaja

Lipanj 30, 2018, Zagreb

Za logističku regresiju za treniranje smo koristili 70% cijelog skupa, što nam ostavlja 30% skupa za testiranje.

Prije provođenja logističke regresije, korištenjem StandardScaler-a značajke su standardizirane. Zbog upotrebe nebalansiranog seta korišten je parametar `class_weight` postavljen na 'balanced' koji automatski daje težinu klasama inverzno proporcionalnu njihovoj frekvenciji. Tablica 2. Prikazuje dobivenu matricu konfuzije.

Actual\Predicted	Normal	Abnormal
Normal	498	253
Abnormal	48	147

Tablica 2. Matrica konfuzije za 20 značajki

Dakle osjetljivost (eng. recall) je 75.38% dok je specifičnost (eng. specificity) 66.31% što daje srednju vrijednost 70.85%.

S obzirom da se neke značajke ne razlikuju toliko za normalne i abnormalne zapise, pokušali smo reducirati broj značajki na 12 koristeći „recursive feature elimination“. Dobiveni ranking značajki redom kojim smo ih naveli je: [1 6 1 5 2 1 1 4 1 3 1 1 8 1 1 1 7 9 1 1]. Koristeći samo tih 12 značajki, nakon standardiziranja i korištenja parametra 'imbalanced', dobiveni su rezultati opisani u matrici konfuzije u Tablici 3.

Actual\Predicted	Normal	Abnormal
Normal	505	246
Abnormal	46	149

Tablica 3. Matrica konfuzije za 12 značajki

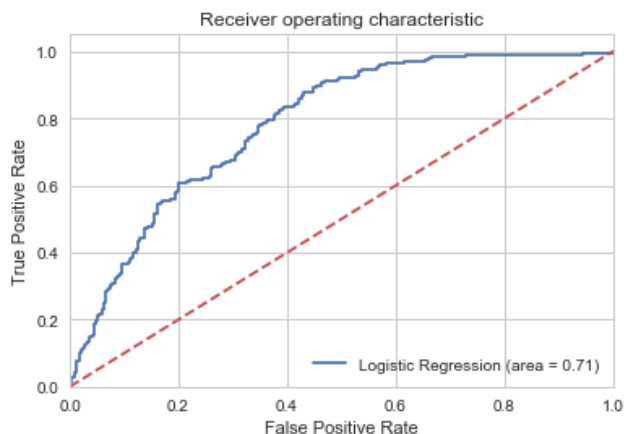
Izračunavanjem se dobije da je osjetljivost 76.41% dok je specifičnost 67.24%, što daje srednju vrijednost 71.83%. Dakle možemo primijetiti da je rezultat dobiven s reduciranim brojem značajki bolji za gotovo 1%.

## 4.4. Rezultati

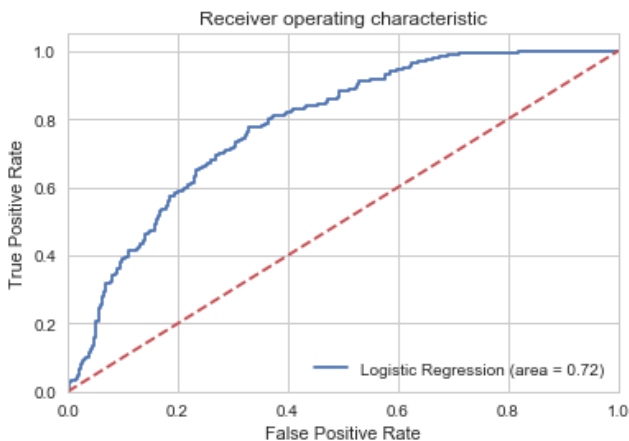
S obzirom da smo koristili poprilično jednostavnu metodu na nebalansiranom setu te samo uzeli u obzir vremenske značajke, rezultat je zadovoljavajući, pogotovo što je osjetljivost veća od specifičnosti, što name je svakako pogodno za medicinski problem.

Ipak, morali smo pribjeći drugim metodama te većem broju značajki kako bismo postigli bolji rezultat.

Za usporedbu, na slikama 9. i 10. prikazane su ROC krivulje za obe veličine skupa značajki.



Slika 9. ROC krivulja za 20 značajki (AUC = 0.71)



Slika 10. ROC krivulja za 12 značajki (AUC = 0.72)

## 5. Klasifikacija koristeći duboku konvolucijsku neuronsku mrežu i MFCC koeficijente

U ovom poglavlju ćemo problem klasifikacije pokušati riješiti po uzoru na model opisan u [4]. Navedeni model je postigao uspješnost od 0.8415 na Physionet 2016 challenge-u.

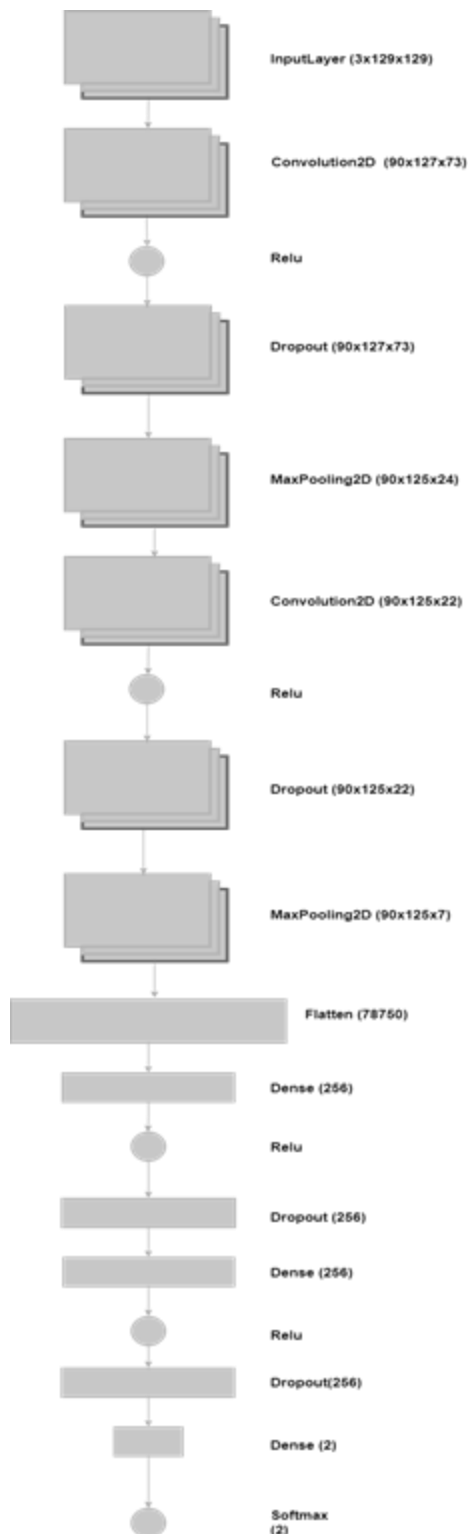
### 5.1. Priprema skupa za učenje modela

Na skupu podataka je prvo proveden undersampling da bi se dobio donekle balansirani skup. Nakon toga je skup podijeljen na skup za treniranje i testni skup u omjeru 85:15. Za potrebe učenja modela za svaki zvučni signal su izračunati MFSC- Mel-Frequency spectral coefficients te

Lipanj 30, 2018, Zagreb



delta i delta-delta značajke a zatim su zbog različite duljine datoteka podijeljene na okvire veličine 128x128. Podaci su zatim normalizirani tako da poprimaju vrijednosti u intervalu [1e-5,1]. Nakon pripreme značajki podaci su promiješani.



Slika 11. Arhitektura konvolucijske neuronske mreže

## 5.2. Model

Za problem klasifikacije srčanog ritma korištena je konvolucijska neuronska mreža. Model je implementiran u programskom jeziku Python, primjenom Keras paketa te je za backend korišten Tensorflow. Arhitektura mreže opisana je na slici 11. Ukupan broj parametara koji je potrebno naučiti je 23 038 432.

Potrebno je odabrati optimalne hiperparametre modela. Zbog činjenice da je učenje modela dugotrajan proces prilagođavati ćemo sljedeće parametre koji će poprimiti vrijednosti na odabranim skupovima:

*Learning rate optimizacijskog algoritma* - 0.0001, 0.00001 i 0.000001

*Dropout parametar* – vrijednost u skupu [0.1,0.5]

*Parametar regularizacije* (l2) – 0.001, 0.0001, 0.00001

*Weight initialization* – „orthogonal“ i „uniform“.

Uz različite kombinacije vrijednosti navedenih parametara korišteni su i sljedeći optimizacijski algoritmi : SGD (momentum=0.9), Adam, RmsProp.

Za potrebe evauluacije modela skup za učenje je podijeljen na novi skup za učenje i evaluacijski skup u omjeru 8:2. Dakle, modeli su ućeni na skupu koji sadrži 1740 normalnih i 1558 abnormalnih okvira. Skup za evaluaciju modela sadrži 423 normalnih i 359 abnormalnih za validaciju. Prilikom biranja modela u obzir smo uzeli preciznost modela i vrijednost funkcije gubitka (categorical cross entropy). Ućenje modela je provedeno tijekom 200 epoha.

Na temelju funkcije gubitka i preciznosti na evaluacijskom skupu odabran je model sa sljedećim parametrima:

- Dropout- 0.1,0.2,0.3,0.4
- Learning rate= 0.00001
- Weight optimizer – Adam
- Regularization coefficient – 0.0001
- Weight inizialization – orthogonal

U tablici 4. su navedena 4 najbolja modela.

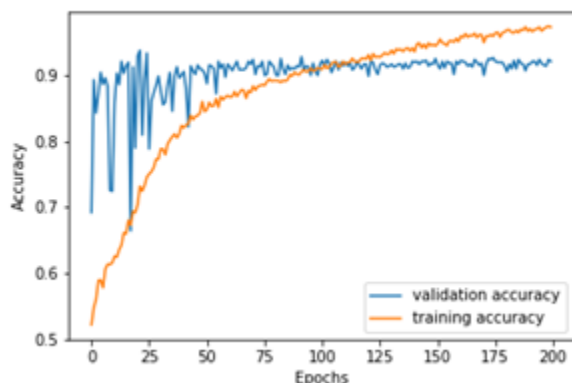
Dropout	l.rate	optimizer	Regularization	init	Loss	Acc
0.1 0.2 0.3 0.4	0.00001	Adam	0.0001	ortho	0.258	0.92
<b>0.1 0.2 0.4 0.4</b>	<b>0.00001</b>	<b>Adam</b>	<b>0.0001</b>	<b>ortho</b>	<b>0.243</b>	<b>0.92</b>
0.25 0.25 0.25 0.25	0.00001	Adam	0.0001	ortho	0.27	0.91
0.2 0.2 0.3 0.3	0.000001	Adam	0.0001	ortho	0.255	0.92

Tablica 4. Najbolji modeli

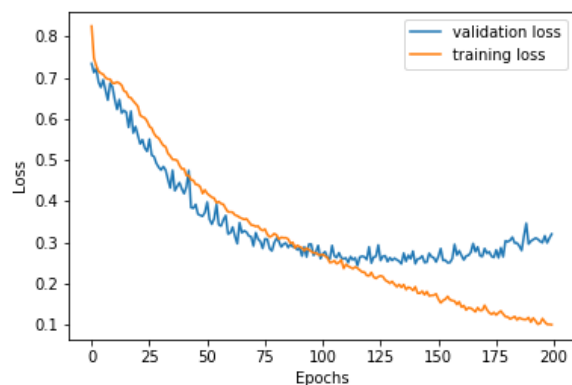
Na slikama 12. i 13. su krivulja ućenja i krivulja gubitka. Moćemo primijetiti da se najmanja vrijednost funkcije gubitka postiće otprilike kada se minimizira razlika između toćnosti na skupu za evaluaciju i toćnosti na skupu za

Lipanj 30, 2018, Zagreb

učenje. Na temelju danih grafova zaključujemo da ne bi trebali imati problema sa overfittingom ili underfittingom modela.



12. Slika krivulje učenje



13. Krivulja funkcije gubitka

### 5.3. Predikcija

Sada kada je dan model koji vrši predikciju na okvirima dobivenim iz zvučnih signala trebamo nekako klasificirati cijeli zvučni signal. Opišimo ukratko kako model klasificira zvučne signale.

Nakon učitavanja signala, računaju se potrebne značajke te se vrši podjela na okvire veličine 128x128. Model zatim klasificira svaki pojedini okvir. Zvučnom signalu se pridjeljuje ona klasa koju je model najčešće predvidio za okvire. Time se modelu daje prostora da pogriješi u klasifikaciji nekih okvira ali da ipak točno klasificira cijeli signal. Takav pristup zapravo pomaže rješavanju problema pogrešne klasifikacije signala zbog raznih šumova u modelu koji otežavaju klasifikaciju.

## 5.4. Rezultati

Nakon odabira parametara klasificiraju se podaci iz test seta. U test setu se nalazi 131 abnormalnih i 136 normalnih zapisa. Rezultati su dani tablicom konfuzije (Tablica 5.). Dakle, osjetljivost je 0.778 a specifičnost 0.867 pa je mjera uspješnosti algoritma 0.8225.

Actual\Predicted	Normal	Abnormal
Normal	118	18
Abnormal	29	102

Tablica 5. Matrica konfuzije

## 5.5. Moguća poboljšanja

Nabrojimo neke metode koji bi mogle utjecati na poboljšanje modela. Primijetimo da je klasifikacija pojedinih signala otežana zbog šumova. Prije prvotne klasifikacije problemu bi se moglo pristupiti tako da se signali prvo klasificiraju po tome da li su dobri za daljnju klasifikaciju po normalnosti i abnormalnosti. Nadalje, kod samog skupa podataka mogli smo isprobati razne metode rješavanja problema nebalansiranosti, na primjer umjesto undersamplinga primijeniti neku metodu oversamplinga čime bi povećali skup sa učenje što bi moglo značajno utjecati na uspješnost modela.

Osim toga, treniranje svakog modela provedeno je kroz 200 epoha za razliku od 660 epoha u [4] gdje je u konačnici naučen i uspješniji model uz istu arhitekturu mreže. Nadalje, uz daljnje eksperimentiranje na većem skupu hiperparametara, to jest primjenom Grid ili Randomized Searcha su isto tako mogući bolji rezultati. Napomenimo da su zadnja dva prijedloga poboljšanja u našem slučaju bila neizvediva zbog dugotrajnog učenja modela, odnosno hardverske ograničenosti.

## 6. Zaključak

Veliki problem kojim smo se susreli je vrlo nebalansiran skup podataka te šumovitost snimaka. Iako dva vrlo različita pristupa, logistička regresija i konvolucijska neuronska mreža dale su dobre rezultate za osjetljivost, dok je konvolucijska mreža uvelike pretekla logističku regresiju što se tiče specifičnosti.

Očekivani score u projektnom prijedlogu name je bio 0.8, što smo i uspjeli ostvariti. Vjerujemo da bi korištenjem svih abnormalnih snimaka umjesto dijela njih (jer nije objavljen test set) uspjeli postići bolje rezultate.

Lipanj 30, 2018, Zagreb

Sljedeći pokušaj bi nam svakako bila implementacija povratne neuronske mreže za koju vjerujemo da bi dala najbolje rezultate. Također smatramo da bi kombinacija vremenskih i frekvencijskih značajki poboljšala rezultate.

## 7. Reference

- [1] <http://www.physionet.org/challenge/2016/> , Classification of normal/abnormal heart sound recordings: the physionet/computing in cardiology challenge 2016
- [2] Singh P. P., Rani P., *An approach to Extract Feature using MFCC*, IOSR Journal of Engineering, Vol. 04, Issue 08
- [3] Springer D.B., Tarassenko L., Clifford G.D. 2014. "Support vector machine hidden semi-Markov model-based heart sound

segmentation", *Computing in cardiology*, (Cambridge, MA: IEEE) pp 625-8

- [4] Maknickas, Vyintas & Maknickas, Algirdas. (2017). Recognition of normal-abnormal phonocardiographic signals using deep convolutional neural networks and mel-frequency spectral coefficients. *Physiological Measurement*. 38. 10.1088/1361-6579/aa7841.

## Autori

**Josip Kiralj**, student, Matematika i statistika,  
Prirodoslovno-matematički fakultet, Zagreb  
**Dajana Jerončić**, student, Računarstvo i matematika,  
Prirodoslovno-matematički fakultet, Zagreb