

# Classification of Heart Sound Recordings

---

## - Projektni prijedlog za kolegij Strojno učenje -

Josip Kiralj, Dajana Jerončić  
Prirodoslovno-matematički fakultet, Zagreb  
30. travnja, 2018.

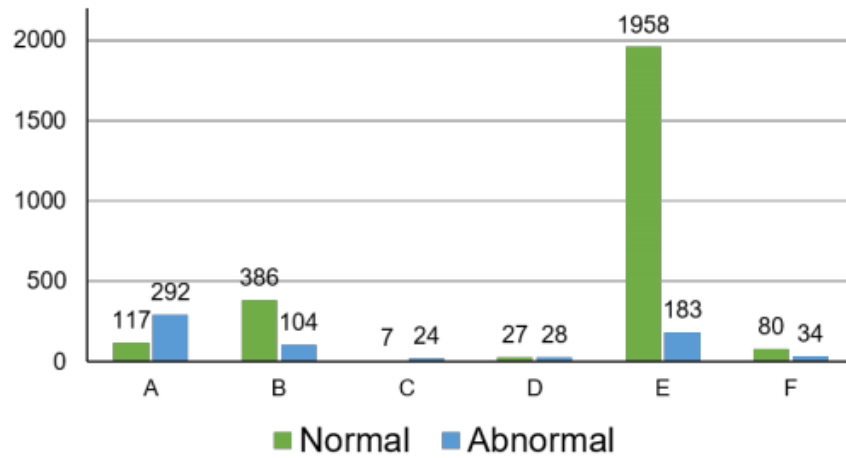
### *1. Uvodni opis problema*

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.

Navedeni problem predstavljen je na web stranici <https://www.physionet.org/challenge/2016/> s koje smo i preuzeli skup podataka koji ćemo koristiti pri klasifikaciji. Natjecanje je završeno 26.08.2016. Problem je od posebnog značaja jer su podaci prikupljeni u stvarnim situacijama. Nerijetko sadržavaju šumove te ostale pozadinske zvukove koji uvelike otežavaju klasifikaciju, ali upravo zbog toga uspješna klasifikacija značila bi korak bliže k pravovremenom dijagnosticiranju kardiovaskularnih bolesti.

#### *1.1 Skup podataka*

Snimke 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. 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 1. Broj snimaka normalnih i abnormalnih otkucaja po dijelovima A-F

Slika 1. prikazuje odnos snimaka normalnih i abnormalnih otkucaja, koji je otprilike u omjeru 4:1, što pokazuje da je skup podataka nebalansiran (o čemu će više govora biti u metodologiji). Također svaki dio dataset-a sadrži i txt datoteku koja ima 3 stupca: prvi stupac je naziv wav file-a, drugi stupac iznosi 1 ako se radi o abnormalnom zapisu, -1 o normalnom, dok zadnji stupac iznosi 1 ukoliko je snimka klasificirana kao „clear“, odnosno 0 ukoliko je „noisy“. Naime, neke snimke su nekvalitetne te sadrže mnoge zvukove iz ambijenta koji otežavaju klasifikaciju te su zbog toga posebno označeni (mnogi natjecatelji su ih izuzeli iz training dataset-a zbog nemogućnosti klasifikacije).

## 2. Cilj i hipoteze istraživanja problema

Cilj istraživanja je pokušati sa što većom točnošću klasificirati snimke srčanog ritma na normalne i abnormalne. Mjera uspješnosti koju ćemo koristiti bit će jednaka onoj s natjecanja. Slika 2. pokazuje značenje oznaka korištenih u formulama (1) i (2) za izračun modificirane osjetljivosti te specifičnosti redom. (Rezultat „unsure“ je opcionalan, te su ga neki natjecatelji koristili kada zbog prenejasne snimke nisu bili u mogućnosti klasificirati snimku.)

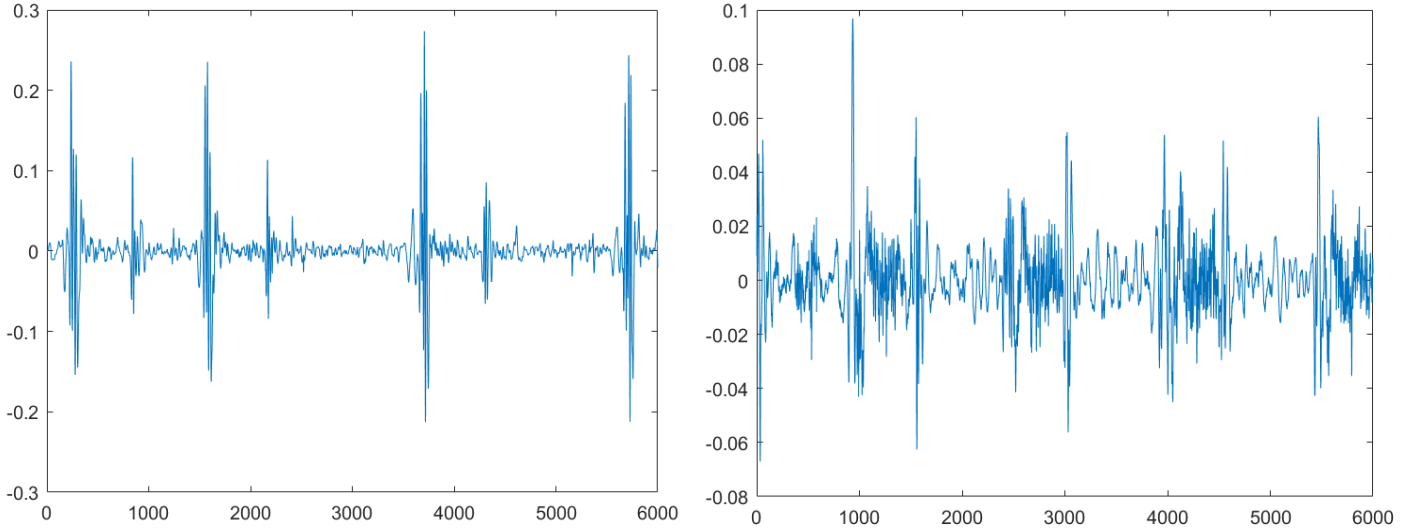
Dijagnoza	Kvaliteta signala	Postotak snimaka	Dobiveni rezultati		
			Abnormal	Unsure	Normal
<b>Abnormal</b>	1	$wa_1$	$Aa_1$	$Aq_1$	$An_1$
<b>(-1)</b>	0	$wa_2$	$Aa_2$	$Aq_1$	$An_2$
<b>Normal</b>	1	$wn_1$	$Na_1$	$Nq_1$	$Nn_1$
<b>(1)</b>	0	$wn_2$	$Na_2$	$Nq_2$	$Nn_2$

Slika 2. Oznake korištene u formulama

$$Se = \frac{wa_1 \times Aa_1}{Aa_1 + Aq_1 + An_1} + \frac{wa_2 \times (Aa_2 + Aq_2)}{Aa_2 + Aq_2 + An_2} \quad (1)$$

$$Sp = \frac{wn_1 \times Nn_1}{Na_1 + Nq_1 + Nn_1} + \frac{wn_2 \times (Nn_2 + Nq_2)}{Na_2 + Nq_2 + Nn_2} \quad (2)$$

Konačna mjera uspješnosti dana je izrazom  $Macc = Se + Sp/2$ , odnosno prosjek je vrijednosti za modificiranu osjetljivost te specifičnost. Ono što ćemo sebi postaviti kao cilj je score iznosa 0.80. S obzirom da su u člancima spomenutim u sljedećem poglavlju koristili jednak način procjene uspješnosti, neće biti teško usporediti dobivene rezultate s postojećima.



Slika 3. Prikaz zvučnih valova kod normalnih otkucaja (lijevo) te abnormalnih (desno)

Pretpostavljamo da će rezultati klasifikacije opravdati predikciju kako će snimke otkucaja koje imaju više šuma te nepravilnije oscilacije zvučnih valova biti klasificirane kao abnormalne. Uzrok tim šumovima kod abnormalnih otkucaja može biti turbulentni, ubrzani protok krvi, dok nepravilnim oscilacijama može biti neujednačeni ritam otkucaja srca. Kao primjer razlike u snimci normalnog i abnormalnog srčanog ritma je slika 3. koja prikazuje zvučne valove (prikazane u MATLAB-u) za normalne otkucaje (lijevo), te abnormalne (desno).

### 3. Pregled dosadašnjih istraživanja

U sklopu *Physionet Classification of Normal/Abnormal Heartsound Recordings* natjecanja korišteni su brojni pristupi, no mi ćemo opisati najuspješnije. Na natjecanju je najveća preciznost/uspješnost dobivena korištenjem ansambla *AdaBoosta* i *konvolucijskih neuronskih mreža*, pristupom koji je opisan u [\[1\]](#). Prije analize, svaki PCG signal je reduciran na 1000Hz. Potom je proveden proces segmentacije koji određuje vremenske periode prvog srčanog tona (S1), drugog srčanog tona (S2), sistole i diastole. Za treniranje *AdaBoost* algoritma korišteno je ukupno 59 značajki koje su podijeljene na dva tipa: *vremenske* i *frekvencijske*. Za treniranje CNN korištene su značajke dobivene dekompozicijom PCG signala na 4 frekvencijska pojasa koja su potom segmentirana. U konačnici, metoda klasificira zapis kao

abnormalan ako je barem jedan klasifikator prepoznao zapisa kao abnormalan. Uspješnost danog algoritma je 0.8602.

Rad [2] koristi duboke *konvolucijske neuronske mreže*. Za razliku od prvog rada, nakon segmentacije signala autori su se oslanjali na MFCC analizu kraćih segmenata PCG signala od kojih svaki počinje otkucajem srca. Značajke dobivene analizom su potom korištene za treniranje DCNN. Iako je uspješnost niža nego u prvom radu (0.83), valja napomenuti da je postignuta najveća specifičnost (0.9521) u odnosu na ostale natjecatelje. Napomenimo da je u oba rada moguća prepreka u klasifikaciji nebalansiranost skupa podataka, problem kojim se nisu pozabavili ni jedni ni drugi.

Oba pristupa u početku segmentiraju PCG signal. Alternativni pristup klasifikaciji je bez segmentacije signala čime se smanjuje računaska složenost. Rad [3] koristi *ansambl* neuronskih mreža kojim osim klasifikacije zvučnih zapisa na normalne/abnormalne klasificira zapise i obzirom na njihovu kvalitetu. Time se pokušava eliminirati pogreška klasifikacije nastala zbog zapisa koji sadrži previše šuma. Isto tako, problem nebalansiranosti pokušao se riješiti *bootstrap* metodom reuzorkovanja. Nedostatak tog rješenja je mogućnost overfittinga. U konačnici, [3] je ostvario drugi najbolji rezultat natjecanja (0.8590). Najveći problem opisanih radova je problem nebalansiranosti podataka i loša kvaliteta pojedinih zapisa.

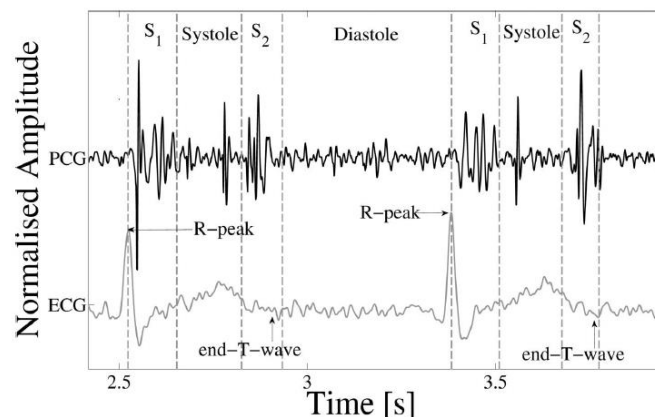
Rad [4], koji se pokazao daleko uspješnijim od svih prijašnjih metoda, objavljen je 2018., dakle nakon završetka natjecanja. Koristi povratne neuronske mreže. Kao i u prije opisanim pristupima, signal prolazi proces segmentacije te se slično kao u [2] PCG signal dijeli na kraće instance određene početkom svakog otkucaja srca. Za određivanje značajki koristi se MFCC analiza. Uspješnost pristupa je 0.9763 što je značajno veće od svih ostalih pristupa.

## 4. Materijali, metodologija i plan istraživanja

Kako je skup podataka nebalansiran, problem ćemo probati riješiti undersamplingom. Napomenimo da je od nekoliko istraženih radova, samo jedan rješavao problem nebalansiranog skupa. Nadalje, skup za testiranje nije objavljen, zbog toga ćemo dobiveni skup podijeliti na training set i test set u omjeru 80:20, što je u broju snimaka 755:189. Međutim, ukoliko rezultati ne budu zadovoljavajući zbog velikog broja neiskorištenih podataka, pribjeći ćemo metodama koje ne mijenjaju veličinu skupa podataka, tzv. „cost sensitive“ tehnikama, koje povećavaju pristranost prema manjim klasama. To se može postići, primjerice većim penaliziranjem pogrešne klasifikacije elemenata iz manje klase tijekom treninga kako bi se poticalo bolje učenje manjih klasa.

Nakon definiranja training i test skupa, potrebno je odrediti značajke. U tu svrhu, za svaki element ćemo provesti segmentaciju (Slika 4.) Springerovom metodom [5] i MFCC analizu te odrediti značajke potrebne za učenje modela. Neke od značajki koje su se u člancima pokazale kao ključne, a za koje smatramo da će biti korisne u našem radu (osim onih dobivenim MFCC analizom) su srednja vrijednost te standardna devijacija za duljine intervala (kao što su duljina intervala između dva otkucaja srca (RR interval), interval od S1, sistole, S2, diastole, omjer trajanja sistole i diastole u odnosu na RR te omjer trajanja sistole u odnosu na diastolu) te za amplitude (npr. omjer srednje amplitude tijekom sistole i one tijekom perioda S1). S obzirom da su se neuronske mreže pokazale kao najefektivniji model, fokusirat ćemo se na njegovu primjenu. Konkretnije, implementirat ćemo povratne neuronske mreže i *feedforward* neuronske

mreže te usporediti rezultate međusobno, ali i s ostalim postojećim rezultatima. Neki od paketa koje ćemo koristiti su: numpy, pandas, sklearn, scipy, matplotlib, imbalanced-learn[6].



Slika 4. Prikaz zvučnog zapisa podjeljenog na 4 stanja: S1, sistola, S2 dijastola

## 5. Očekivani rezultati predloženog projekta

Kao što smo već spomenuli, očekivana mjera uspješnosti predloženog projekta je 0.8 te smatramo da ćemo visokim score-om uspjeti približiti se najboljim postojećim rješenjima. Nadamo se da ćemo uspjeti riješiti problem nebalansiranih podataka za razliku od ostalih kandidata. Također očekivamo da će povratna neuronska mreža dati bolje rezultate od feedforward neuronske mreže, s obzirom da su dosadašnja istraživanja pokazala kako je upravo to najefikasnija metoda za ovaj problem. I na kraju, očekujemo da će klasifikacija zvučnih zapisa označenih kao „noisy“ biti relativno uspješna u odnosu na ostale natjecatelje te da ćemo tako uspjeti klasificirati i one snimke koje nisu snimane u kliničkim okruženjima, što bi značilo mogućnost neinvazivnog pravovremenog dijagnosticiranja srčanih anomalija.

## 6. Literatura

- [1] Cristhian Potes, Saman Parvaneh, Asif Rahman, Bryan Conroy, [\*Ensemble of Feature-based and Deep learning-based Classifiers for Detection of Abnormal Heart Sounds\*](#)
- [2] Jonathan Rubin, Rui Abreu, Anurag Ganguli, Saigopal Nelaturi, Ion Matei, Kumar Sricharan, [\*Classifying Heart Sound Recordings using Deep Convolutional Neural Networks and Mel-Frequency Cepstral Coefficients\*](#)
- [3] Morteza Zabihi, Ali Bahrami Rad, Serkan Kiranyaz, Moncef Gabbouj, Aggelos K. Katsaggelos, [\*Heart Sound Anomaly and Quality Detection using Ensemble of Neural Networks without Segmentation\*](#)
- [4] Siddique Latif, Muhammad Usman, Junaid Qadir, Rajib Rana, *Abnormal Heartbeat Detection Using Recurrent Neural Networks*
- [5] Springer D. B., Tarassenko L. and Clifford G. D. 2014, *Support vector machine hidden semi-Markov model-based heart sound segmentation. In: Computing in Cardiology*, (Cambridge, MA: IEEE) pp 625-8
- [6] Lemaitre G., Nogueira F., Aridas C. K., *Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning*