

# Predikcija VTVF stanja

Istraživanje podataka u bioinformatiki

Profesor: Gordana Pavlović - Lažetić

Student: Jelena Jeremić, 1099/2021

Matematički fakultet

# Sadržaj

- 1 Zadatak
- 2 Podaci
- 3 Priprema podataka
- 4 Modeli klasifikacije
- 5 Rezultati
- 6 Zaključak

# Kratak pregled

- 1 Zadatak
- 2 Podaci
- 3 Priprema podataka
- 4 Modeli klasifikacije
- 5 Rezultati
- 6 Zaključak

# Cilj i opis rada

- Naprasna srčana (SCD) smrt jedan je od najčešćih uzroka smrti i dešava se u roku kraćem od sat vremena od pojave prvih simptoma. Prevencija se ugradnjom defibrilatora (ICD) koji u odgovarajućim trenucima šalje električne impulse kako bi reaktivirao srčani rad
- Ventrikularna fibrilacija (VF) se dešava kada srce radi brzim, izuzetno nepravilnim ritmom, zbog potpunog poremećaja sprovođenja srčanih impulsa. Često se ventrikularna fibrilacija javlja nakon pojave ventrikularne tahikardije (VT) (veoma brz rad srca), a često i sama ventrikularna tahikardija ukazuje na ozbiljnije poremećaje, poput lošeg snabdevanja srca krvlju, ili na stanje posle infarkta

# Cilj i opis rada

- Većina ljudi gubi svest za nekoliko sekundi i zahtevaju hitnu medicinsku pomoć. Bez reanimacije i vraćanja normalnog srčanog ritma, smrt nastupa za nekoliko minuta
- Cilj rada je pronaći parametre date kroz kliničku sliku pacijenta kako bi što sigurnije mogli predvideti VTVF stanje koje prethodi SCD-u

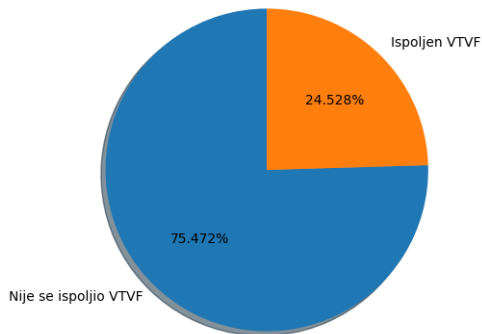
# Kratak pregled

- 1 Zadatak
- 2 Podaci
- 3 Priprema podataka
- 4 Modeli klasifikacije
- 5 Rezultati
- 6 Zaključak

# Podaci

- Dati su podaci 106 pacijenata. Od toga, kod 80 njih nije detektovano VT/VF stanje
- Atributi: vreme\_do\_prve\_VT, vreme\_do\_prve ICD\_terapije i uzrok\_smrti su imali preko 70% nedostajućih vrednosti

# Podaci



Slika1: Udeo pojave VTFV stanja



# Kratak pregled

- 1 Zadatak
- 2 Podaci
- 3 Priprema podataka**
- 4 Modeli klasifikacije
- 5 Rezultati
- 6 Zaključak

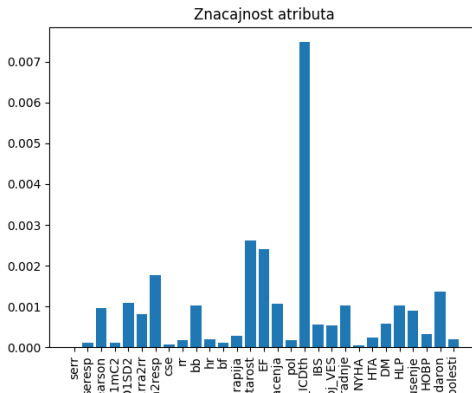
# Priprema podataka

Radi lakšeg rada sa podacima prilikom klasifikacije, nominalni atributi su prebačeni u diskretne. Podaci za klasifikaciju su podeljeni na dva skupa.

Prvi skup - podaci iz kojih su uklonjeni atributi koji su meusobno visoko korelisani i oni koji nisu od značaja za predikciju.

Drugi skup - dodatno pročišćen prvi skup, uklonjeni su atributi sa niskom varijansom.

# Priprema podataka

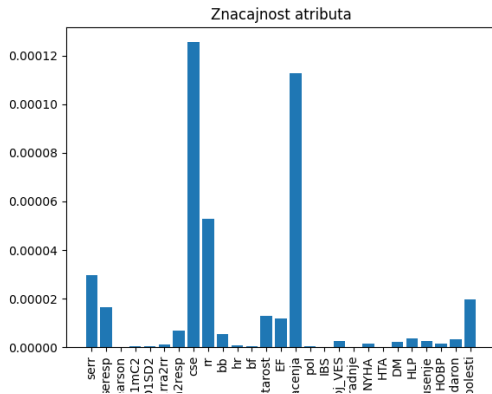


Slika2: Značajnost atributa

# Priprema podataka

Dodatno, izdvojena su još dva skupa podataka: - bez atributa  
dužina\_praćenja - bez atributa ICDterapija i vreme\_do\_prve\_ICDth

# Priprema podataka

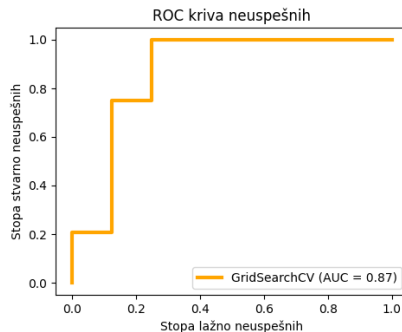
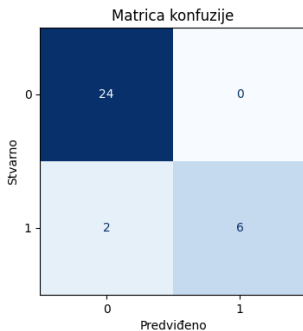


**Slika3:** Značajnost atributa nakon izbacivanja atributa koji sadrže podatke o ICD terapiji

# Kratak pregled

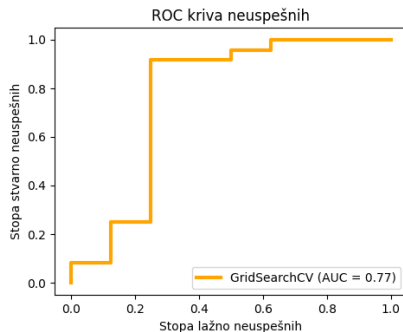
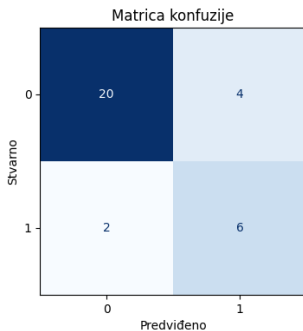
- 1 Zadatak
- 2 Podaci
- 3 Priprema podataka
- 4 Modeli klasifikacije**
- 5 Rezultati
- 6 Zaključak

# Neuronska mreža



Slik4: Model sa zadržanim atributom dužina praćenja

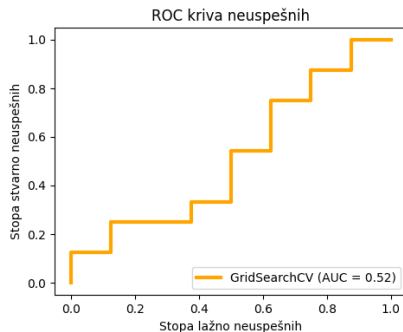
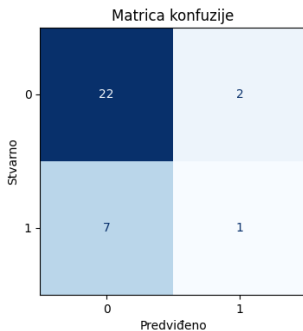
# Neuronska mreža



Slika5: Model bez atributa dužina praćenja

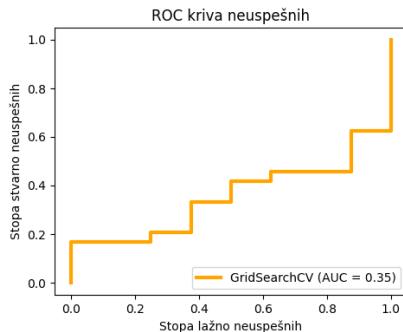
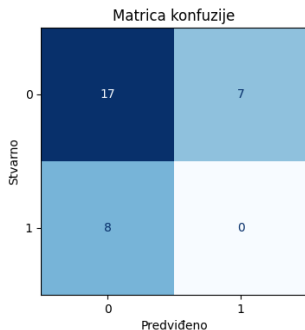


# Neuronska mreža



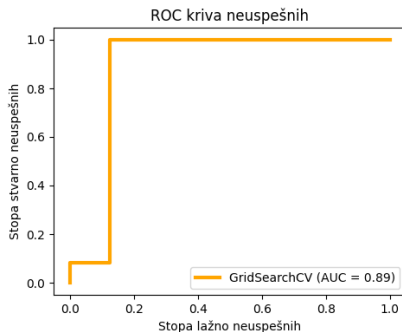
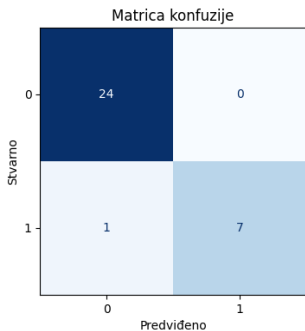
Slika6: Model bez atributa koji sadrže informacije o ICD terapiji

# Neuronska mreža



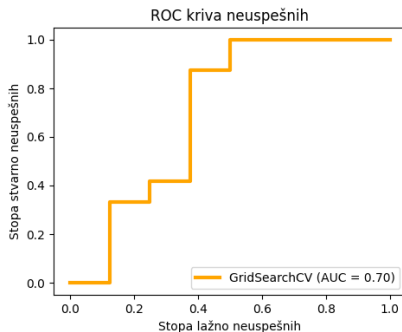
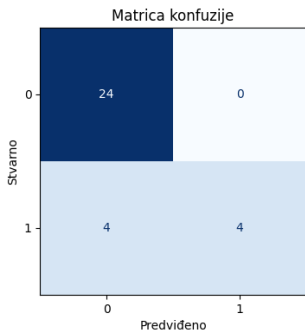
**Slika7:** Model bez atributa koji sadrže informacije o ICD terapiji i bez dužine praćenja

# Metod potpornih vektora



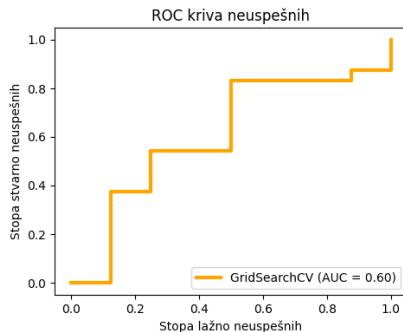
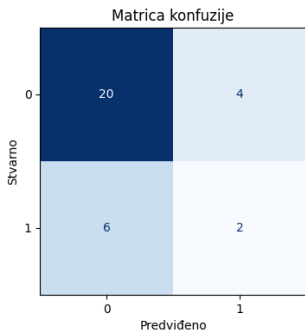
Slika8: Model sa zadržanim atributom dužina praćenja

# Metod potpornih vektora



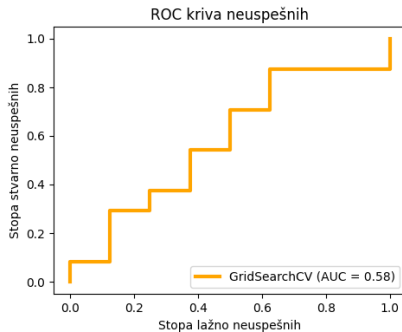
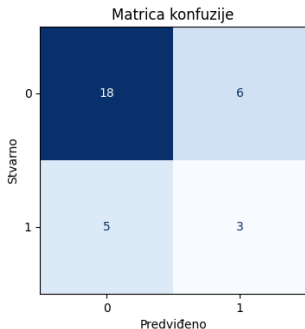
Slika9: Model bez atributa dužina praćenja

# Metod potpornih vektora



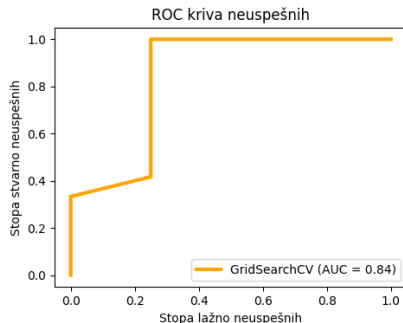
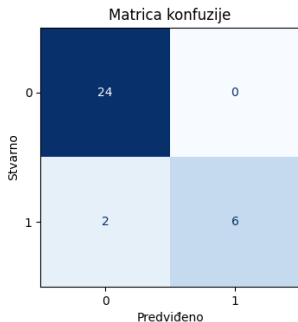
Slika10: Model bez atributa koji sadrže informacije o ICD terapiji

# Metod potpornih vektora



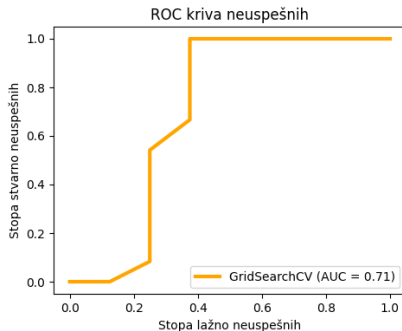
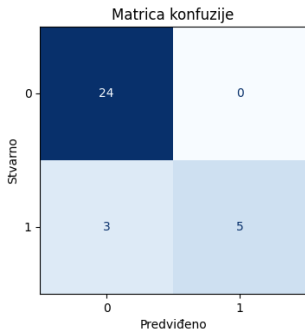
**Slika11:** Model bez atributa koji sadrže informacije o ICD terapiji i bez dužine praćenja

# Slučajna šuma



Slika12: Model bez atributa dužina praćenja

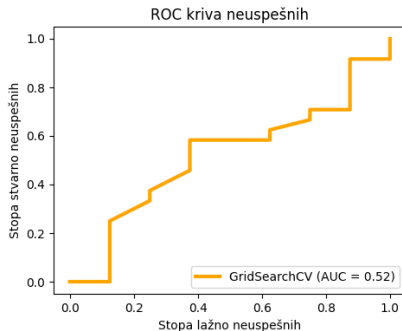
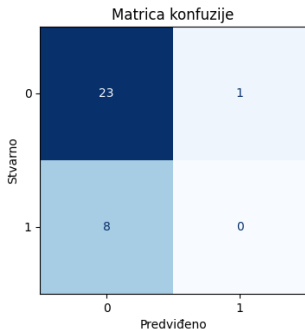
# Slučajna šuma



Slika13: Model bez atributa dužina praćenja

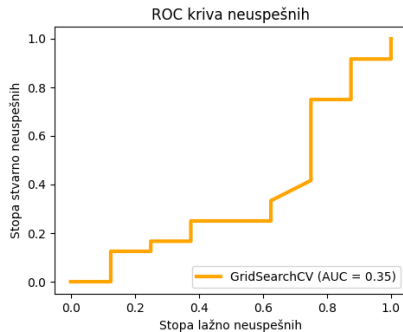
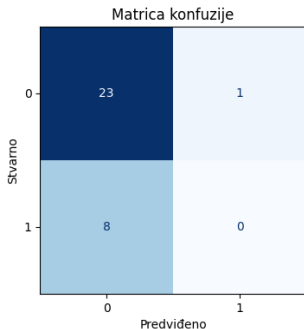


# Slučajna šuma



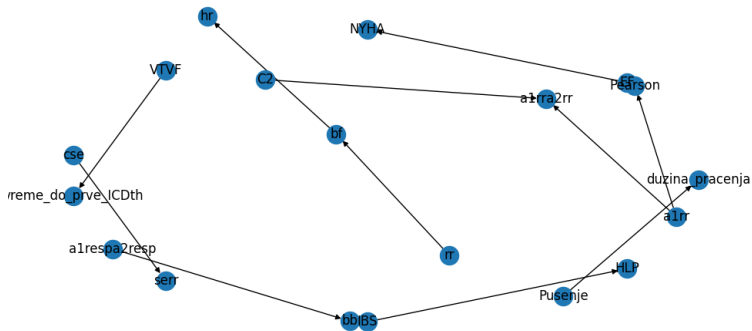
Slika14: Model bez atributa koji sadrže informacije o ICD terapiji

# Slučajna šuma



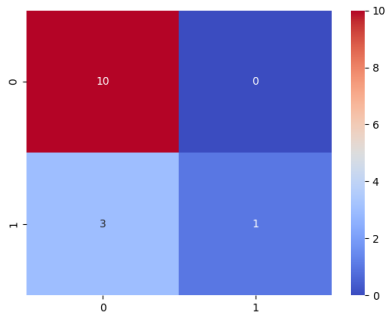
**Slika15:** Model bez atributa koji sadrže informacije o ICD terapiji i bez dužine praćenja

# Bajesova mreža



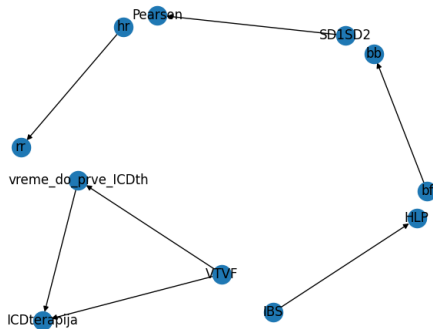
Slika16: PC estimator, korelacija zasnovana na Pearson-u

# Bajesova mreža



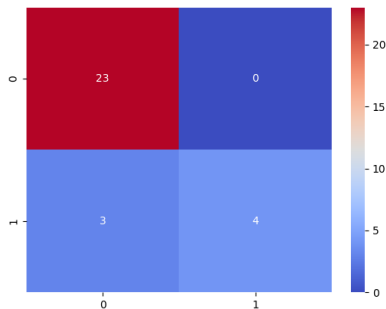
Slika17: PC estimator, matrica konfuzije

# Bajesova mreža



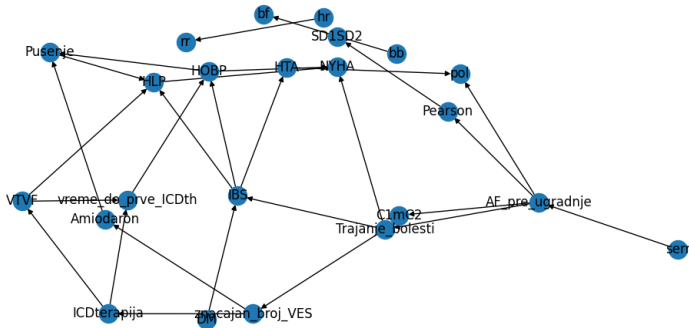
Slika18: PC estimator, funkcija pogotka K2

# Bajesova mreža



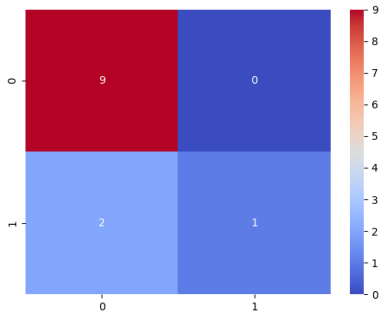
Slika19: PC estimator, funkcija pogotka K2

# Bajesova mreža



Slika20: HillClimber estimator

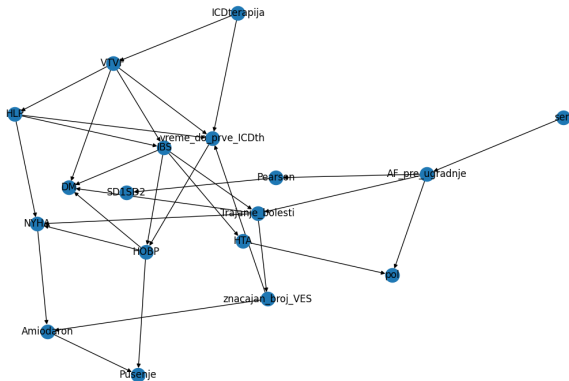
# Bajesova mreža



Slika21: HillClimber estimator

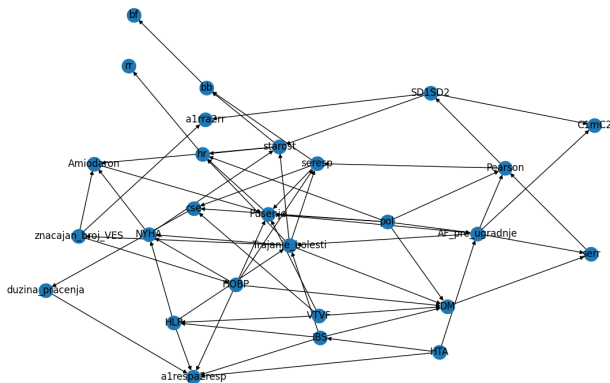


# Bajesova mreža



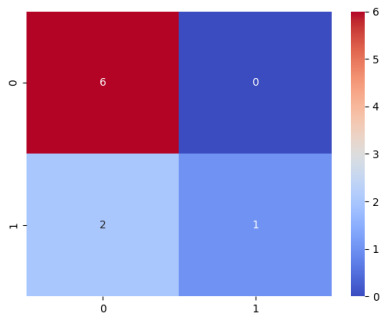
Slika22: HillClimber estimator, funkcija pogotka BDeuScore

# Bajesova mreža



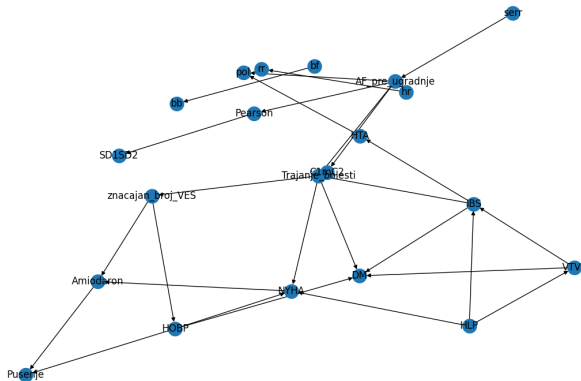
**Slika23:** HillClimber estimator, bez atributa koji sadrže informacije o ICD terapiji

# Bajesova mreža



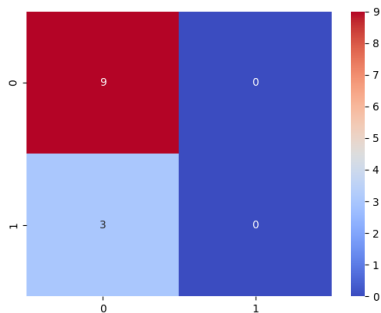
**Slika24:** HillClimber estimator, bez atributa koji sadrže informacije o ICD terapiji

# Bajesova mreža



**Slika25:** HillClimber estimator, bez atributa koji sadrže informacije o ICD terapiji i bez dužine praćenja

# Bajesova mreža



**Slika26:** HillClimber estimator, bez atributa koji sadrže informacije o ICD terapiji i bez dužine praćenja

# Kratak pregled

- 1 Zadatak
- 2 Podaci
- 3 Priprema podataka
- 4 Modeli klasifikacije
- 5 Rezultati**
- 6 Zaključak

# Rezultati i zaključak

Atributi vreme\_do\_prve\_ICDth, ICDTerapija i dužina\_praćenja su visoko korelisani sa ciljnim atributom. Uklanjanjem ovih atributa, preciznost modela je u svakom algoritmu klasifikacije bila u nekoj meri narušena.

Kod Bajesa, uklanjanjem ova tri atributa, modeli su uspjeli da pronau vezu izmeu ciljnog atributa i atributa HLP, cse, IBS, EM i DM.

# Kratak pregled

- 1 Zadatak
- 2 Podaci
- 3 Priprema podataka
- 4 Modeli klasifikacije
- 5 Rezultati
- 6 Zaključak**



# Rezultati i zaključak

Kao najbolji modeli pokazali su se metod potpornih vektora i neuronske mreže. Kao najgori pokazao se Bajesov model. Korišćenje većeg skupa podataka bi verovatno doprinelo pronalaženju boljeg modela, zbog postojanja varijansi u rezultatima.