

**Ispitna pitanja iz Mašinskog učenja 1**  
**Studijski program: Biomedicinsko inženjerstvo**

2024/2025.

**1. PREGLED LINEARNE ALGEBRE**

- Iz ove oblasti ne javljaju se klasična teorijska pitanja, ali se traži razumevanje i primenljivo znanje sledećih koncepata:
  - Množenje matrica i množenje vektora.
  - Norma vektora. Ortogonalnost vektora i ortogonalna projekcija.
  - Linearna nezavisnost vektora. Rang matrice.
  - Inverzija matrice. Ortogonalnost i ortonormalnost matrice.
  - Vektorski prostori i linearne transformacije.
  - Determinanta matrice.
  - Karakteristični vektori, karakteristične vrednosti i njihova interpretacija.

**2. PREGLED VEROVATNOĆE I STATISTIKE**

- Iz ove oblasti ne javljaju se klasična teorijska pitanja, ali se traži razumevanje i primenljivo znanje sledećih koncepata:
  - Osnovne osobine verovatnoće.
  - Uslovna verovatnoća.
  - Teorema totalne verovatnoće i Bayesova teorema.
  - Binarna klasifikacija i brojni pokazatelji njene uspešnosti.
  - Slučajne promenljive i gustina raspodele verovatnoće.
  - Matematičko očekivanje, varijansa i standardna devijacija.
  - Slučajni vektori. Vektor srednjih vrednosti i kovarijansna matrica. (Ne)korelisanost i (ne)zavisnost slučajnih promenljivih.
  - Gaussova raspodela. Centralna granična teorema.

**3. UVOD U MAŠINSKO UČENJE**

- Objasniti osnovne pojmove iz oblasti mašinskog učenja – uzorak, obeležje, prostor obeležja, model.
- Kada je, za određeni problem, određeni vektor obeležja dobar, a kada loš? Koja svojstva određeni vektor obeležja može, a ne mora imati?
- Objasniti pojmove generalizacije i natprilagođenja, uz navođenje odgovarajućih primera<sup>1</sup>.
- Navesti i objasniti uloge pojedinih komponenata sistema za mašinsko učenje.
- Navesti i objasniti razlike između pojedinih tipova predikcionih problema, uz navođenje odgovarajućih primera.

---

<sup>1</sup> Kod ovog, kao i kod ostalih ispitnih pitanja kod kojih se očekuje navođenje konkretnih primera, od studenata se očekuje da, ako je moguće, navedu originalne primere.

- Objasniti razlike između ranog i kasnog učenja, uz navođenje odgovarajućih primera.
- Objasniti pojam normalizacije podataka i navesti dva najčešća načina na koje se ona vrši.
- Objasniti razvojni ciklus sistema za mašinsko učenje.

#### 4. VALIDACIONI POSTUPCI

- Kako se u praksi vrši izbor optimalnog modela (izbor vrednosti određenih slobodnih parametara) i kako se vrši procena stope njegove greške? Objasniti pojam stvarne stope greške.
- Objasniti u čemu se sastoji procena greške određenog modela resupstitucijom i navesti osnovne mane ovakvog načina procene greške.
- U čemu su prednosti metoda procene greške modela ponovnim uzorkovanjem u odnosu na procenu greške na nezavisnom skupu za testiranje? U čemu su osnovne razlike između metoda unakrsne validacije i *bootstrap* metoda? Koje su prednosti *bootstrap* metoda?
- Koji se načini ponovnog uzorkovanja koriste kod metoda unakrsne validacije i koje su prednosti i mane pojedinih načina u odnosu na ostale?
- Objasniti osnovni princip i razloge za korišćenje trostruke podele podataka.

#### 5. LINEARNA I LOGISTIČKA REGRESIJA

- Formulirati problem regresije i objasniti ga na odabranom primeru. Kada se regresija smatra linearnom?
- Objasniti princip linearne regresije za slučaj jednog obeležja. Za dati skup za obuku  $\{x^{(i)}, y^{(i)}\}$ ,  $i = 1, 2, \dots, N$ , koja su dva osnovna načina za nalaženje vrednosti odgovarajućih parametara  $\vartheta_0$  i  $\vartheta_1$ ?
- Objasniti osnovne principe, prednosti i nedostatke metode gradijentnog silaska:
  - a) kao opšte metode za nalaženje minimuma funkcije više promenljivih;
  - b) u posebnom slučaju kada se primenjuje za rešavanje problema linearne regresije.
- Objasniti ulogu parametra brzine učenja kod metode gradijentnog silaska. Da li je kod ove metode neophodno aktivno prilagođavati brzinu učenja u zavisnosti od približavanja lokalnom minimumu i zbog čega? Kojih se još korisnih saveta treba pridržavati kod primene ove metode?
- Objasniti osnovni princip analitičke minimizacije funkcije cene u rešavanju problema linearne regresije. Kada ima smisla primenjivati ovaj pristup a kada ne?
- Kakvi se problemi javljaju ako se pri analitičkoj minimizaciji funkcije cene u rešavanju problema linearne regresije ispostavi da među obeležjima za obuku ima međusobno izuzetno korelisanih? Navesti jedan način za prevazilaženje ovih problema.
- Objasniti pojam regularizacije u kontekstu linearne regresije i navesti dva često korišćena načina na koji se ona može izvršiti.
- Objasniti princip, prednosti i mane uvođenja baznih funkcija kod linearne regresije. Dati definiciju i ilustraciju sigmoida.
- Objasniti osnovni princip logističke regresije. Šta u ovom slučaju predstavljaju vrednosti diskriminantnih funkcija pojedinih klasa?

#### 6. BAYESOVA TEORIJA ODLUČIVANJA

- Objasniti primenu Bayesove teoreme u kontekstu problema klasifikacije. Šta je apriorna verovatnoća klase, šta aposteriora verovatnoća klase a šta izglednost konkretne opservacije u

određenoj klasi? Zbog čega se u praksi retko pristupa problemima odlučivanja direktnom primenom ovih koncepata?

- Izvesti formulaciju LR testa (testa na osnovu količnika izglednosti) za slučaj dve klase različitih apriornih verovatnoća, uz pretpostavku da se uzorak klasifikuje u onu klasu za koju je odgovarajuća aposteriorna verovatnoća veća.
- Objasniti pojam Bayesovog rizika i dati njegovu matematičku definiciju.
- Izvesti pravilo odlučivanja koje minimizuje Bayesov rizik za slučaj dve klase, za date apriorne verovatnoće klasa i date vrednosti cena pogrešne/ispravne klasifikacije.
- Ukratko opisati kako se pravilo odlučivanja:
  - a) na osnovu minimizacije verovatnoće greške,
  - b) na osnovu minimizacije Bayesovog rizika,uopštava na slučaj više od dve klase.
- Definirati pojam diskriminantne funkcije i navesti čemu je jednaka diskriminantna funkcija određene klase u slučaju primene Bayesovog kriterijuma, odlučivanja na osnovu aposteriorne verovatnoće (MAP) i odlučivanja na osnovu izglednosti (ML).

## 7. KVADRATNI KLASIFIKATORI

- Navesti tipične primere izgleda 1-D i 2-D Gaussovih raspodela i na njima objasniti parametre koji ih opisuju.
- Izvesti opšti oblik diskriminantne funkcije za klasu čija je apriorna verovatnoća  $P(\omega_i)$  a gustina raspodele verovatnoće  $\mathcal{N}(\mu, \Sigma)$ . Zbog čega se klasifikator na bazi ovakve diskriminantne funkcije naziva kvadratnim klasifikatorom?
- Neka je potrebno formulisati pravilo odlučivanja u problemu klasifikacije pri čemu su gustine raspodele verovatnoće svih klasa Gaussove raspodele. Kada će granice odlučivanja između klasa biti linearne (hiperravni), a kada ne?

## 8. PARAMETARSKA ESTIMACIJA GUSTINE RASPODELE VEROVATNOĆE

- Navesti osnovna dva pristupa koja se koriste za estimaciju (uslovnih) gustina raspodela verovatnoće uzoraka pojedinih klasa.
- Navesti osnovne razlike između estimacije na osnovu maksimizacije izglednosti skupa uzoraka za obuku i Bayesove estimacije.
- Objasniti princip estimacije nepoznatog vektora parametara  $\theta$  određene raspodele na osnovu maksimizacije izglednosti skupa uzoraka za obuku.
- Objasniti pojmove pristrasnosti (necentriranosti) i varijanse procene određene slučajne veličine.
- Izvesti izraz za procenu srednje vrednosti  $\mu$  1-D Gaussove raspodele na osnovu maksimizacije izglednosti skupa uzoraka za obuku, pod uslovom da je varijansa ove raspodele poznata i da iznosi  $\sigma^2$ .
- Izvesti izraze za procenu srednje vrednosti  $\mu$  i varijanse  $\sigma^2$  1-D Gaussove raspodele na osnovu maksimizacije izglednosti skupa uzoraka za obuku. Zbog čega procena  $\sigma^2$  dobijena na ovaj način nije centrirana i koja se procena može koristiti umesto nje?
- Objasniti osnovni princip Bayesove estimacije. Ako je potrebno estimirati nepoznatu srednju vrednost  $\mu$  Gaussove raspodele  $\mathcal{N}(\mu, \sigma^2)$  čija je varijansa  $\sigma^2$  poznata, komentarisati (date) izraze za srednju vrednost i varijansu ove procene:

$$\mu_N = \frac{N\sigma_0^2}{N\sigma_0^2 + \sigma^2} \bar{x}_N + \frac{\sigma^2}{N\sigma_0^2 + \sigma^2} \mu_0 \quad \sigma_N^2 = \frac{\sigma_0^2 \sigma^2}{N\sigma_0^2 + \sigma^2} = \frac{1}{N/\sigma^2 + 1/\sigma_0^2}$$

pri čemu su  $\mu_0$  i  $\sigma_0^2$  srednja vrednost i varijansa priora, a  $N$  je veličina skupa uzoraka za obuku.

## 9. ESTIMACIJA GUSTINE RASPODELE VEROVATNOĆE POMOĆU KERNELA (KDE)

- Objasniti osnovni princip, prednosti i nedostatke histograma.
- Dati opštu formulaciju neparametarske estimacije gustine raspodele verovatnoće.
- Koja dva načina estimacije gustine raspodele verovatnoće prozilaze iz opšte formulacije neparametarske estimacije? Šta je kod ova dva načina isto, a šta različito?
- Objasniti pojam Parzenovog prozora, dati njegovu definiciju i objasniti njegove nedostatke.
- U kakvoj su matematičkoj vezi stvarna gustina raspodele verovatnoće i njena KDE estimacija? Kako bi kernel trebalo da izgleda pa da očekivanje KDE estimacije bude upravo jednako stvarnoj gustini raspodele verovatnoće? Zbog čega to u praksi ipak nije dobro rešenje?
- Objasniti prednosti korišćenja glatkih kernela. Zbog čega je izbor širine kernela u praksi problematičan i koji se koraci mogu preduzeti u pravcu rešavanja tog problema?
- Zbog čega primena standardnog izraza za procenu gustine raspodele verovatnoće na osnovu kernela u višedimenzionalnom prostoru:

$$\hat{p}_{KDE}(\mathbf{x}) = \frac{1}{N \cdot h^D} \sum_{n=1}^N K\left(\frac{\mathbf{x} - \mathbf{x}^{(n)}}{h}\right)$$

često nije dobro rešenje? Šta se dobija primenom separabilnih kernela?

- Objasniti osnovnu ideju i prednosti naivnog Bayesovog klasifikatora.

## 10. ESTIMACIJA GRV I KLASIFIKACIJA METODOM $k$ NAJBЛИŽIH SUSEDA ( $kNN$ )

- Objasniti osnovni princip estimacije gustine raspodele verovatnoće metodom  $k$  najbližih suseda i navesti osnovne nedostatke ove metode.
- Pokazati na koji način se metodom  $k$  najbližih suseda može aproksimirati (optimalan) Bayesov klasifikator.
- Objasniti uticaj izbora parametra  $k$  na tačnost estimacije gustine raspodele verovatnoće metodom  $k$  najbližih suseda.
- Koji se problemi mogu javiti ako pojedina obeležja u okviru primene metode  $k$  najbližih suseda imaju bitno različite dinamičke opsege (redove veličine) i šta se može preduzeti u cilju rešavanja tog problema?
- Navesti i ukratko objasniti dva najčešće korišćena načina na koja se može ubrzati pretraga prostora obeležja u cilju identifikacije  $k$  najbližih suseda datoј tački.

## 11. STABLA ODLUKE

- Ilustrovati upotrebu stabla odluke u rešavanju problema klasifikacije/regresije na odabranom primeru.
- Objasniti proceduru obuke (formiranja) stabla odluke u slučaju klasifikacije/regresije.
- Diskutovati računarsku kompleksnost klasifikacije/regresije primenom stabla odluke, i to posebno u fazi obuke a posebno u fazi primene.
- Objasniti prednosti i mane stabala odluke u rešavanju problema klasifikacije/regresije.

## 12. LINEARNA KLASIFIKACIJA I PERCEPTRONSKO UČENJE

- Dati geometrijsku interpretaciju linearne (binarne) klasifikacije. Kako se ovaj pristup generalizuje na slučaj više od dve klase?
- Definisati pojam i matematički model perceptrona. Kako se pomoću ovakvih struktura može rešavati problem klasifikacije u dve klase, a kako problem klasifikacije u više od dve klase?
- Opisati metod perceptronskog učenja. U čemu je razlika između grupnog i pojedinačnog režima perceptronskog učenja? Kada algoritam perceptronskog učenja konvergira ka ispravnom rešenju, a kada ne?

## 13. NEURONSKE MREŽE

- Kako se perceptron (neuron) može modifikovati tako da vrši nelinearnu klasifikaciju i koja su ograničenja tog pristupa?
- Dati primer jednostavne neuronske mreže s jednim skrivenim slojem i predstaviti zavisnost aktivacionih funkcija od ulaznih veličina uz korišćenje matričnog zapisa.
- Navesti osnovne principe obuke neuronske mreže. Kada se smatra da je u postupku obuke nastupila konvergencija? Da li postoji rizik od upadanja algoritma za obuku u lokalni minimum?
- Kojih se korisnih saveta treba pridržavati u postupku obuke neuronskih mreža? Kako se vrši inicijalizacija sinaptičkih težina i zbog čega baš tako? Šta treba uraditi sa ulaznim obeležjima u pripremnoj fazi?
- Od čega zavisi brzina konvergencije algoritma za obuku neuronske mreže i kako se na nju može uticati?
- Koji problemi nastaju ako je struktura mreže složenija (npr. ako mreža ima više slojeva ili više neurona po slojevima) nego što je to za dati skup za obuku potrebno? Kako je moguće izbeći preteranu obučenost mreže (eng. *overtraining*)?

## 14. METODE NA BAZI VEKTORA NOSAČA

- Objasniti osnovnu ideju i dati matematičku formulaciju optimizacionog problema kod klasifikatora maksimalne margine. Objasniti pojam vektora nosača. Šta su glavni nedostaci ovakvog klasifikatora?
- Objasniti osnovnu ideju i dati matematičku formulaciju optimizacionog problema kod klasifikatora blage margine. Šta su u ovom slučaju vektori nosači? Šta su ograničenja ovakvog klasifikatora?
- Objasniti osnovnu ideju mašina na bazi vektora nosača. Objasniti kako se konstrukcija nelinearne hiperpovršni razdvajanja u niskodimenzionalnom prostoru obeležja svodi na konstrukciju (linearne) hiperravni razdvajanja u visokodimenzionalnom prostoru obeležja. Objasniti ulogu kernela u ovom postupku.
- Navesti prednosti metode vektora nosača u odnosu na neuronske mreže u rešavanju problema klasifikacije.

## 15. KLASSTERIZACIJA

- Objasniti pojam klasterizacije, navesti osnovne pristupe klasterizaciji i glavne razloge zbog kojih se ona vrši.
- Koja tri koraka obuhvata rešavanje problema neparametarske klasterizacije?
- Koje se mere bliskosti između uzoraka standardno koriste u statističkoj klasterizaciji?

- Koje se kriterijumske funkcije standardno koriste u statističkoj klasterizaciji?
- Koji se algoritmi za minimizaciju/maksimizaciju kriterijumske funkcije standardno koriste u statističkoj klasterizaciji? Po čemu su oni međusobno slični, a po čemu su različiti?
- Objasniti (na odabranom primeru) funkcionisanje algoritma  $k$  srednjih vrednosti (eng.  $k$ -means) i navesti njegove osnovne nedostatke.
- Objasniti osnovni princip ISODATA algoritma. Po čemu je on sličan  $k$ -means algoritmu, a po čemu je različit od njega?
- Koja su dva osnovna pristupa hijerarhijskoj klasterizaciji? Koje se mere rastojanja između klastera koriste kod algoritama zasnovanih na povezivanju i koje su osnovne razlike između njih?
- Pokazati funkcionisanje algoritma hijerarhijske klasterizacije zasnovane na minimalnom/maksimalnom/srednjem rastojanju između klastera (na odabranom primeru).

## 16. RAZLAGANJE NA GLAVNE KOMPONENTE (PCA) I LINEARNA DISKRIMINANTNA ANALIZA (LDA)

- Objasniti potrebu za smanjenjem dimenzionalnosti i dva osnovna pristupa smanjenju dimenzionalnosti. Ako se transformacijom polaznog vektora obeležja  $\mathbf{x}$  dobija ciljni vektor obeležja  $\mathbf{y}$  niže dimenzionalnosti, kada je ta transformacija optimalna?
- Neka se izdvajanje obeležja u cilju redukcije dimenzionalnosti vrši na osnovu funkcije  $\mathbf{y} = f(\mathbf{x})$ . Koja se dva osnovna pristupa razlikuju u zavisnosti od toga koje kriterijume koristi funkcija cilja koju treba maksimizovati/minimizovati?
- Neka je, u okviru PCA, potrebno predstaviti slučajni vektor  $\mathbf{x}$  u prostoru sa manjim brojem dimenzija. Da bi on pri tome bio predstavljen sa minimalnom srednjekvadratnom greškom, kako treba odabrati vektore baze  $\phi_i$  u novom vektorskom prostoru? Šta je glavni nedostatak ovakve reprezentacije?
- Šta je osnovni zadatak LDA? Ako se polazni vektorski prostor projektuje na jedan pravac (na vektorski prostor dimenzije 1), navesti funkciju cilja koja se pri tome koristi i objasniti zbog čega je takav izbor opravdan, a zatim diskutovati i krajnji izraz za optimalni pravac  $\mathbf{w}^*$  na koji je potrebno izvršiti projekciju.
- Na koje se sve načine primena LDA u binarnoj klasifikaciji može uopštiti na slučaj više od dve klase? Koja je najveća dimenzionalnost novog prostora obeležja koji se primenom LDA može konstruisati? Koja su druga važna ograničenja LDA?

Pitanja na ispitu odnose se na gradivo pokriveno pitanjima sa ovog spiska. Formulacija pitanja na ispitu može biti u manjoj ili većoj meri izmenjena u odnosu na formulaciju navedenu ovde, a mogu se javiti i pitanja koja ovde nisu navedena, kao i jednostavniji računski zadaci (zbog čega kod sebe treba imati džepni kalkulator). Na pitanja treba davati što direktnije i što potpunije odgovore i pre svega treba pokazati razumevanje gradiva. Ukoliko se slajdovi koriste kao pomoćni materijal pri pripremi ispita, treba izbegavati da se odgovor na pitanje svede isključivo na sadržaj slajda.