



Domaći zadaci iz Prepoznavanja oblika 2022/2023

Zadatak 1.

Izabrati jednu od dve ponuđene opcije za izradu ovog domaćeg zadatka.

Opcija 1.

Na sajtu predmeta* dostupna je baza zabeleženih brzina po x i y osi, kao i pritisaka zabeleženih prilikom ispisivanja slova na grafičkoj tabli pod nazivom *PO_slova_brzine* (u njoj se nalazi fajl 'PO_slova.mat'). Za jedno slovo prva vrsta su brzine po x osi u ekvidistantnim trenucima, druga vrsta brzine po y osi, a treća vrsta pritisak (sve u ekvidistantnim trenucima). Za datu bazu projektovati inovativni sistem za prepoznavanje 10 slova po izboru zasnovan na testiranju hipoteza. Zbog malog broja dostupnih odbiraka za svako slovo, nije potrebna podela na trening i test skup. Uzeti približno jednak broj odbiraka za svako od slova. Nije dozvoljeno koristiti obeležja korišćena na času vežbi ($\max(v_{x/y})$ - $\min(v_{x/y})$), kao i $\max(v_{x/y})$ i $\min(v_{x/y})$.

- Za jedan primerak svakog od odabranih slova prikazati njegov oblik (poziciju na y osi u zavisnosti od pozicije na x) i njegove brzine po x i y osi. Prokomentarisati povezanost brzine po x i y osi sa načinom ispisivanja slova.
- Rezultate klasifikacije prikazati u obliku matrice konfuzije.
- Opisati projektovani sistem, obrazložiti izbor obeležja i prikazati i prokomentarisati karakteristične primere pravilno i nepravilno klasifikovanih slova (ukoliko ih ima).
- Odabrati dva slova i dva obeležja takva da su odabrana dva slova separabilna u tom prostoru.
- Za slova i obeležja pod c) projektovati parametarski klasifikator po izboru i iscrtati klasifikacionu liniju.

Opcija 2. (3 bonus poena)

U timu predmeta na MSTeams platformi dostupna je baza slikanih šaka koje pokazuju „papir“, „kamen“ i „makaze“. Projektovati inovativni sistem za prepoznavanje pokazanih znakova zasnovan na testiranju hipoteza.



- Detaljno opisati algoritam za obradu slike i odabir obeležja koji prethodi samoj klasifikaciji. Algoritam treba da bude što robusniji (na različite osvetljaje, položaje šaka, načine pokazivanja znakova itd).
- Izvršiti podelu na trening i test skup. Rezultate klasifikacije test skupa prikazati u obliku matrice konfuzije.
- Odabrati dva znaka i dva obeležja takva da su odabrani znakovi što separabilniji u tom prostoru. Prikazati histogram obeležja za oba slova i prokomentarisati njihov oblik.
- Za slova i obeležja pod c) projektovati parametarski klasifikator po izboru i iscrtati klasifikacionu liniju.

Zadatak 2.

Generisati po $N = 500$ odbiraka iz dveju dvodimenzionih bimodalnih klasa:

$$\Omega_1 \sim P_{11} \cdot N(M_{11}, \Sigma_{11}) + P_{12} \cdot N(M_{12}, \Sigma_{12}),$$

$$\Omega_2 \sim P_{21} \cdot N(M_{21}, \Sigma_{21}) + P_{22} \cdot N(M_{22}, \Sigma_{22}).$$

Parametre klasa samostalno izabrati.

- Na dijagramu prikazati odbirke.
- Iscrtati kako teorijski izgledaju funkcije gustine verovatnoće za raspodele klasa i uporediti ih sa histogramom generisanih odbiraka.
- Projektovati *Bajesov* klasifikator minimalne greške i na dijagramu, zajedno sa odbircima, skicirati klasifikacionu liniju. Uporediti grešku klasifikacije konkretnih odbiraka sa teorijskom greškom klasifikacije prve i druge vrste za datu postavku.
- Projektovati klasifikator minimalne cene tako da se više penalizuje pogrešna klasifikacija odbiraka iz prve klase.
- Ponoviti prethodnu tačku za *Neuman-Pearson*-ov klasifikator. Obrazložiti izbor $\varepsilon_2 = \varepsilon_0$.
- Za klase oblika generisanih u prethodnim tačkama, projektovati *Wald*-ov sekvencijalni test pa skicirati zavisnost broja potrebnih odbiraka od usvojene verovatnoće grešaka prvog, odnosno drugog tipa.

Zadatak 3.

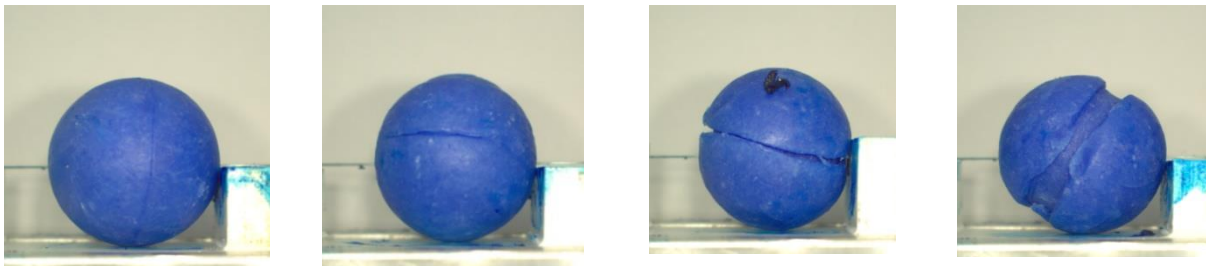
Izabrati jednu od dve ponuđene opcije za izradu ovog domaćeg zadatka.

Opcija 1.

- Generisati **tri klase** dvodimenzionalnih oblika. Izabrati funkciju gustine verovatnoće oblika tako da klase budu linearno separabilne.
 - Za tako generisane oblike izvršiti projektovanje linearnog klasifikatora jednom od tri iterativne procedure. Rezultate prikazati u obliku matrice konfuzije. Detaljno opisati postupak klasifikacije.
 - Ponoviti prethodni postupak korišćenjem metode željenog izlaza. Analizirati uticaj elemenata u matrici željenih izlaza na konačnu formu linearnog klasifikatora.
- Generisati **dve klase** dvodimenzionalnih oblika koje jesu separabilne, ali ne linearno, pa isprojektovati kvadratni klasifikator metodom po želji.

Opcija 2. (3 bonus poena)

Za bazu slika sanitarnih kuglica, koja je dostupna na sajtu predmeta* isprojektovati inovativni sistem za detekciju postojanja kuglice na slici i, u slučaju da kuglica postoji, detekciju kuglica sa defektnom konturom (oblikom u ravni slike). Rezultate prikazati u obliku matrice konfuzije. Izveštaj treba da sadrži kratki opis isprojektovanog sistema, obrazložen izbor obeležja, kao i karakteristične primere pravilnog i nepravilnog rada klasifikatora. Primetiti da kuglica može na sebi imati i manju pukotinu, a da ona po obliku (konturi) bude regularna (druga kuglica u nizu na Slici 1). Primeri kuglica ispravne i neispravne konture su dati na slici.



Slika 1. Primeri slika sa regularnom konturom (levo) i defektnom konturom (desno).

Zadatak 4.

1. Generisati po $N = 500$ dvodimenzionih odbiraka iz četiri klase koje će biti linearno separabilne. Preporuka je da to budu Gausovski raspodeljeni dvodimenzioni oblici. Izabrati jednu od metoda za klasterizaciju (c mean metod, metod kvadratne dekompozicije) i primeniti je na formirane uzorke klasa. Izvršiti analizu osetljivosti izabranog algoritma na početnu klasterizaciju kao i srednji broj potrebnih iteracija. Takođe izvršiti analize slučaja kada se apriorno ne poznaje broj klasa.
2. Na odbircima iz prethodne tačke izabrati jednu od metoda klasterizacije (metod maksimalne verodostojnosti ili metod grana i granica) i primeniti je na formirane uzorke klasa. Izvršiti analizu osetljivosti izabranog algoritma na početnu klasterizaciju kao i srednji broj potrebnih iteracija. Takođe izvršiti analize slučaja kada se apriorno ne poznaje broj klasa.
3. Generisati po $N = 500$ dvodimenzionih odbiraka iz dve klase koje su nelinearno separabilne. Izabrati jednu od metoda za klasterizaciju koje su primenjive za nelinearno separabilne klase (metod kvadratne dekompozicije ili metod maksimalne verodostojnosti) i ponoviti analizu iz prethodnih tačaka.

Napomena. Student treba da pripremi izveštaj koji će sadržati tekst zadataka, kratak teorijski osvrt na rešenje problema, dobijene rezultate i komentare. Izveštaji za sva četiri domaća zadatka treba da budu u jednom dokumentu. Jasno označene kodove smestiti u zaseban folder i postaviti ih kao prilog izveštaju u okviru MSTeams platforme. Domaći zadaci se brane zbirno, a ne pojedinačno.

* (<http://automatika.etf.rs/index.php/sr/prepoznavanje-oblika-os4po/os4po-materijali>)