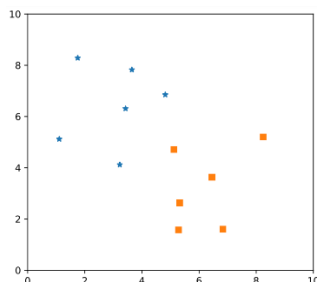
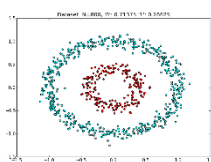
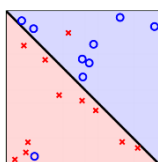


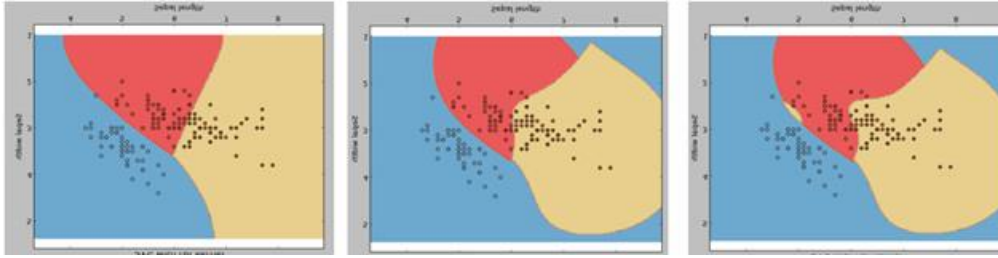
- U SVM (*Support Vector Machine*) algoritmu, dimenzija rezultujuće hiper-ravni zavisi od:
  - broja obeležja
  - broja primera u skupu podataka
  - broja vrednosti koje uzima ciljna varijabla
  - svega navedenog.
- Na slici ispod ilustrujte kako izgleda granica odluke za SVM model. Označite na ovoj slici vektore potpore (*support vectors*) i marginu hiper-ravni.

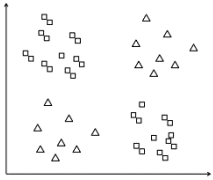


- Tačno ili netačno:
  - U svojoj osnovnoj formulaciji, SVM je primenljiv na probleme višekategorijske klasifikacije.
  - Brisanje (uklanjanje) vektora potpore promeniće poziciju hiper-ravni.
  - Uklanjanje korektno klasifikovane instance iz skupa podataka koja se nalazi daleko od granice odluke promeniće poziciju hiper-ravni.
- Šta su *linearly separable* podaci?
- Kako možemo primeniti SVM model u slučaju kada podaci NISU linearly separable? Za slike ispod označite koji pristup bi bio bolji za taj tip podataka.



- U SVM algoritmu se funkcija koju optimizujemo u cilju pronalaska optimalne hiper-ravni zove:
  - Categorical cross-entropy loss
  - Binary cross-entropy loss
  - Hinge loss
  - Nijedna od navedenih.
 Ilustrujte kako izgleda ova funkcija na grafiku.
- Efektivnost SVM klasifikatora zavisi od:
  - Selekcije kernela
  - Parametara kernela
  - Soft margin parametra C
  - Svega navedenog.
- Recimo da smo iscrtavali vizuelizaciju za različite vrednosti parametra C (*slack penalty*) u *soft margin SVM*:  $\min \frac{1}{2} \theta^T \theta + C \cdot (\text{br. grešaka})$ . Ako C1 odgovara prvom grafiku, C2 drugom, a C3 trećem, šta je tačno od sledećeg:
  - $C1 = C2 = C3$ ;    b.  $C1 > C2 > C3$ ;    c.  $C1 < C2 < C3$ ;    d. Ne možemo reći.



9. Objasnite uticaj parametra  $C$  kod *soft margin* SVM:  $\min \frac{1}{2} \theta^T \theta + C \cdot (\text{br. grešaka})$ . Šta se dešava sa marginom kada je  $C = 0$  i kada je  $C = \infty$ ? Kako u praksi biramo optimalnu vrednost  $C$ ?
10. Za date primere (kvadrati su jedna klasa, a trouglovi druga), koji od sledećih kernela bismo mogli iskoristiti da razdvojimo dve klase:
- 

a. Linearni kernel

b. *Gaussian* RBF kernel

c. Polinomijalni kernel

d. Nijedan od navedenih.
11. Možemo li primeniti SVM na problem višekategorijske klasifikacije? Ako da, na koji način?
12. Dat vam je binarni klasifikacioni problem gde je broj primera u skupu podataka  $N$  manji od broja obeležja  $D$ . Ako biste primenjivali SVM nad ovim skupom podataka, koji od sledećih kernela će najverovatnije biti odgovarajući: a) linearni; b) polinomijalni; c) RBF;
13. Recimo da imamo skup podataka sa  $D=10$  obeležja i  $N=5000$  primera. Nakon što smo trenirali model logističke regresije, vidimo da su i trening i validaciona greška visoke. Koji od sledećih koraka deluju da će rezultovati boljim performansama:
- Uvećati regularizacioni parametar  $\lambda$
  - Trenirati SVM model sa RBF kernelom
  - Kreirati nova polinomijalna obeležja
  - Trenirati SVM model sa linearnim kernelom
  - Smanjiti broj primera u skupu podataka
14. Šta je „kernel trik“?
15. Razmotrite sledeća dva primera (tačke u 2D):  $A = (1, 2)$ ,  $B = (2, 4)$ . Iskoristite sledeće mapiranje u 6-dimenzioni prostor:  $\Phi(x) = (x_1^2, x_2^2, \sqrt{2}x_1x_2, \sqrt{2}x_1, \sqrt{2}x_2, 1)$ :
- $\Phi(A) =$
  - $\Phi(B) =$
- Sada izračunajte skalarni proizvod ovde dve tačke:
- $\Phi(A)^T \Phi(B) =$
- Kernel funkcija koja odgovara transformaciji datoj iznad je polinomijalni kernel drugog stepena:  $K(x, y) = (x^T y + 1)^2$ . Iskoristite ovu funkciju da izračunate:
- $K(A, B) =$