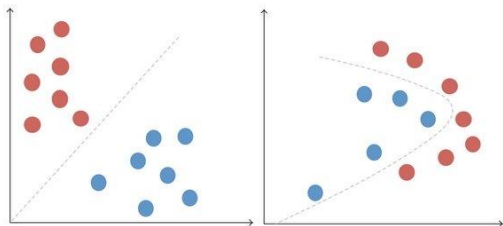
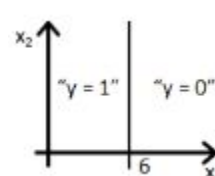
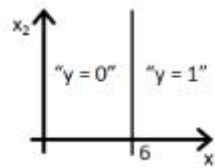
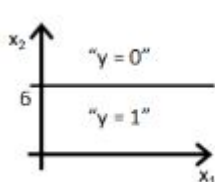
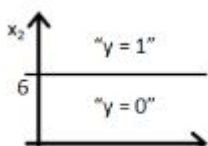


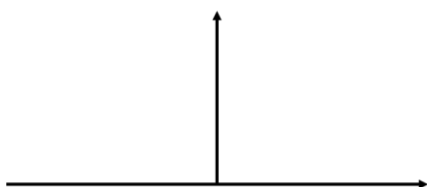
1. Logistička regresija koristi se za (1) regresiju (2) klasifikaciju (3) klasterovanje (4) sve navedeno.
2. Logistička regresija kao izlaz vraća _____. Granica odluke koju proizvodi je linearna/nelinearna (zaokružiti tačno).
3. Koji grafik prikazuje linearnu granicu odluke? Čime je položaj granice odluke određen?



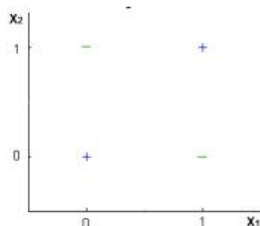
4. Recimo da smo trenirali model logističke regresije $h_{\theta} = g(\theta_0 + \theta_1 x_1 + \theta_2 x_2)$ i dobili da je $\theta_0 = 6$, $\theta_1 = 0$, $\theta_2 = -1$. Koji od sledećih grafika predstavlja granicu odluke dobijenog klasifikatora?



5. Kako bi izgledao grafik ako bismo u prethodnom pitanju zamenili vrednosti x_1 i x_2 ?
6. Zašto se funkcija greške (cost) koju smo koristili za linearnu regresiju ne može koristiti kod logističke regresije?
7. Kojim postupkom optimizujemo parametre kod logističke regresije?
8. Kako izgleda cost funkcija kod logističke regresije? Napišite formulu i ilustrujte je na grafiku (ne zaboravite da obeležite ose i granice na osama).



9. Ako bi kao izlaz modela logističke regresije koristili otežinjenu sumu ulaza (kao u linearnoj regresiji), vrednost izlaza bi se kretala između _____ i _____, a mi želimo da izlaz bude između _____ i _____.
10. Može li model logističke regresije perfektno klasifikovati podatke na slici ispod? Možete koristiti samo varijable x_1 i x_2 .

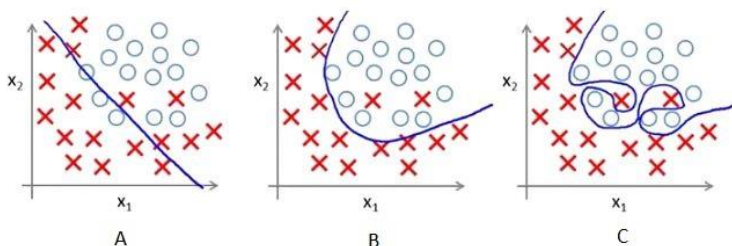


11. Tačno ili netačno:

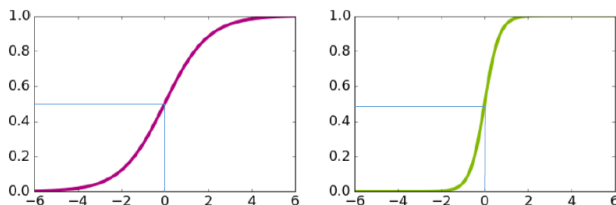
- a. Funkcija greške koju optimizujemo kod logističke regresije je konveksna.
- b. Može nam se desiti da primenom gradijentnog spusta na funkciju greške logističke regresije upadnemo u lokalni optimum. Zato nam je potreban napredniji algoritam za optimizaciju parametara.
- c. Pošto je granica odluke koju proizvodi model logističke regresije linearna, to znači da ne možemo dobiti krivu liniju koja razdvaja klase.
- d. Model logističke regresije (u svojoj osnovnoj formulaciji) možemo primeniti isključivo na problem binarne klasifikacije.
- e. Obavezno moramo normalizovati vrednosti obeležja pre primene modela logističke regresije.
- f. Model logističke regresije može biti *overfit*-ovan.

12. Koje metrike ne možemo upotrebiti za evaluaciju modela logističke regresije? (1) accuracy, (2) F-measure, (3) MSE, (4) R^2 , (5) precision, (6) recall.

13. Koji od sledećih modela je optimalan, koji *underfit*-ovan, a koji *overfit*-ovan?



14. U slučaju *overfitting*-a, vrednosti parametara θ će po magnitudi biti _____ (mali/veliki). Ovo znači da će granica odluke biti sličnija grafiku _____ (levo/desno). Verovatnoće koje model procenjuje na izlazu biće _____ (male/velike).



15. Kako se možemo boriti protiv *overfitting*-a u logističkoj regresiji?

16. Koji je efekat L1 regularizacije kod logističke regresije?

17. Ako imamo C klasa, kod *1-vs-all* pristupa višekategorijskoj klasifikaciji fitujemo _____ modela, a kod *1-vs-1* pristupa _____ modela.

18. Kod pojedinačnog modela *1-vs-all* pristupa, vršimo binarnu klasifikaciju da razdvojimo klasu _____ od klase _____.

19. Kod pojedinačnog modela *1-vs-1* pristupa, vršimo binarnu klasifikaciju da razdvojimo klasu _____ od klase _____.

20. Pristup *1-vs-1* je računarski _____ (zahtevniji/manje zahtevan) u odnosu na *1-vs-all* pristup ako imamo umereno velik skup podataka.

21. U slučaju velikog broja klasa C , *1-vs-all* pristup je problematičan jer _____.

22. *Softmax regression* koristimo za _____. Izlaz ovog modela predstavlja _____, a kao klasifikacionu odluku biramo _____.