

① Uvod i osnovni koncepti

- (a) Za skup označenih primjera $D = \{(x^i, y^i)\}_i = \{((0,0), 1), ((0,1), 1), ((1,1), 0)\}$ razmatramo linearan klasifikacijski model H s parametrima $\theta \in \mathbb{R}^3$. Odredite $|H|$ i $|VS_{H,0}|$ za slučajevе (1) $X = \{0,1\}^2$ (2) $X = \mathbb{Z}^2$ (3) $X = \mathbb{Z}^2$ uz dodatno $((-1,1), 0)$

- (b) Raspodajemo sljedećim primjerima za vježbu:

$$D = \{(x^i, y^i)\}_i = \{((1,1), 0), ((0,2), 0), ((2,3), 0), ((3,1), 1), ((4,3), 1)\}$$

Skicirajte funkcije gubitka L kao funkcije od $y_h(x)$ za gubitak (1) linearne regresije (2) perceptronu (3) logističke regresije (4) stroja P.V. Pozivajući se na skice funkcija gubitaka, skicirajte predviđive hipoteze u ova 4 algoritma. Učinite isto za skup podataka u logii je dodan primjer $((8,1), 1)$

② Regresija

- (a) L2-regularizirana empirijska pogreška lin. modela je $E(w|D) = \frac{1}{2} \sum_{i=1}^N (y^i - \vec{w}^T \vec{x}^i)^2 + \frac{\lambda}{2} \|w\|^2$. Krenuvši od tog izraza, izvedite rješenje u matričnoj formi za težine w^* koje minimiziraju ovu pogrešku. Kako se pri regularizaciji tretira w_0 i zašto?

- (b) Koristimo model lin. reg. s L2 regularizacijom u ulaznom prostoru $X = \mathbb{R}^2$. Model je treniran na D koji sadrži $x^{(1)} = (1, 5)$, $x^{(2)} = (3, 2)$, $x^{(3)} = (2, 4)$, $x^{(4)} = (7, 7)$. Prilikom koristimo prelikovanje $\Phi(x) = (1, x_1, x_2, x_2 - x_1)$. Bi li vektor optimalnih težina w^* bio nestabilan ili neizračunljiv? Ako da, detaljno objasnite vrake i moguće rješenje problema.

③ Linearni diskriminativni modeli i povezani linearni modeli

- (a) Izvedite algoritam perceptrona krenuvši od funkcije koja apsproximira broj vratnih klasifikacija. Konvergira li taj algoritam za svaki skup primjera D ? Zašto?
- (b) $D = \{(x^i, y^i)\}_i = \{((-3,1), 0), ((-3,3), 0), ((1,2), 1), ((2,1), 1), ((1,-2), 1), ((2,-3), 2)\}$. Za klasifikaciju koristimo regresiju i sheme OVR. Definirajte maticu dizajna i vektor ornaka \mathbf{y} za svaki od triju modela te skicirajte (otprilike) granice između klasa (i za binaru i za zajednički model). Koje su prednosti ova naspram OVO?

4) Logistička regresija

(a) Izvedite pogrešku uvažene entropije $E(\hat{y}|D)$ kao negativnu log-izglednost

na skupu označenih primjera, koji je (najviše) interval u kojem se može naći vrijednost pogreške za hipotezu koja na skupu od 5 pozitivnih i 5 negativnih primjera svih 10 primjera klasificira ispravno?

(b) Slicirajte pogreške učenja i ispitivanja $E(\hat{y}|D)$ u ovisnosti o broju iteracija za $\lambda=0$ i $\lambda=100$ i za slučaj linearno odugivih i linearne neodugivih primjera (s grafičnom podajnjem krivulje)

(c) Objasnite kvazi-Newtonov postupak i motivaciju za njegovu primjenu

(d) Izvedite problem vrde maksimalne marge, pa pomoću njega formulirajte problem mjeke maksimalne marge. Je li mješana marga potrebna ako preslikavamo u prostor značajki? Zasto?

(e) Za klasifikaciju primjera u uličnom pr. $X=\mathbb{R}^2$ koristimo SUM s polinomijalnom jezgrnom funkcijom. $K(\vec{x}, \vec{z}) = (\vec{x}^T \vec{z} + 1)^2$. Izvedite priputnicu preslikavanja $\phi(\vec{x})$ za $n=2$. Kada će broj parametara neparametarske inačice ovog modela biti vedi od broja parametara njegove parametarske inačice? (U ovoj slučaju, parametri su vektori realnih brojeva)

(f) Koristimo SUM s Gauss. jezgr. funk. i provodimo optimizaciju hiperparametara C i γ pretraživanjem po rešetci. Okvirno slicirajte očekivane izoblikture pogreške učenja i ispitne pogreške u prostoru $C \times \gamma$ (duće slice). Naznačite područje prenovečnosti i podnovečnosti te detaljno objasnite zašto i na koji način hiperparametri C i γ utječu na složenost modela.

5) Neparametarske metode i ansamblji

(a) Formalno definirajte klasifikacijske modele k -NN i težinski k -NN. Objasnite problem problematike dimenzionalnosti i kako se on manifestira kod k -NN

(b) Objasnite bagging i napišite pseudokod algoritma random forest