IIA ML Esercitazioni – parte 3 (appunti in sintesi di esercizi svolti in aula)

Si leggano le note inziali nell'esercitazione 1.

Di seguito alcune note relative a esercizi svolti nelle lezioni di esercitazione 3:

1- "Due passi nel gradiente"

a) Definire un compito supervisionato di regressione con modello lineare.
b) Vi è fornito il dato (esempio) di training nella forma (input x, output desiderato y) di valori (1,4). Disegnare nel piano cartesiano due possibili soluzioni esatte al problema, assumendo in una delle due che la retta passi dall'origine degli assi. Scrivere l'equazione dei due modelli individuati (mostrando i valori di w_0 e w_1).
c) Si provi l'*algoritmo di discesa del gradiente*: si calcoli $w_0(t)$ e $w_1(t)$ *per* $t=1$ *e* 2 (due passi dell'algoritmo) assumendo *eta* (*learning rate*) pari a ¼, i valori inziali dei pesi $w_0(0)=0$ e $w_1(0)=0$, la seguente *loss* (e mantenendo il 2 nei Δw_i):

$$E(\mathbf{w}) = \sum_{p=1}^{l} (y_p - h(\mathbf{x}_p))^2$$

d) Determinare l'equazione del modello $h(x)$ risultante dal punto c e disegnare la soluzione nel piano cartesiano.

Soluzioni (solo risultato per verifica, svolgere personalmente)

a) Vedere le definizioni a lezione.
Ad esempio: Dati l esempi di training in formato $(\mathbf{x}_p, y_p)$ $p=1..l$, $y_p \in R$, trovare i valori di $\mathbf{w}$ *in* $h(\mathbf{x})= \mathbf{w}^T\mathbf{x}$ tali che $E(\mathbf{w})$ sia minimo (e si approssimi la funzione target). Se c'è una variabile sola (come nel caso del proseguo ai punti successivi) si può specificare $\mathbf{x}= [1, x]$ oppure usare anche semplicemente x usando la forma con w_0 esterno.
Oppure: Dati l esempi di training trovare una buona approssimazione della funzione target a valori reali con $h(\mathbf{x})= \mathbf{w}^T\mathbf{x}$, minimizzando il rischio empirico sulle l coppie note espresso come $E(\mathbf{w})=$ … (e cerca di minimizzare il rischio R).
b) Graficare per esempio le funzioni $h(\mathbf{x}) =4$ ($\mathbf{w}$=(4,0)) e $h(\mathbf{x})=4x+0$ ($\mathbf{w}$=(0,4)) dove $\mathbf{w}=(w_0,w_1)$.
Esistono ovviamente molte altre soluzioni come (anche) $h(\mathbf{x})=x+3$, etc.
c) C'è qui un solo dato $(\mathbf{x}_p, y_p)$ = (1,4), per cui $l=1$ nella sommatoria dei p della regola delta che va qui a semplificarsi, quindi si procede con $\mathbf{w} = 0 + eta\ 2\ (y-h(\mathbf{x}))\ \mathbf{x} = 0 + eta\ 2\ (y-\mathbf{w}^T\mathbf{x}))\ \mathbf{x}$.
Al primo passo si ha $h(x)= 0\ x +0$ *e si procede con* $w_0=0+¼\ 2\ (4-0)=2$, $w_1= 0+¼\ 2\ (4-0)\ 1=2$.
Al secondo si parte da $h(x)=2\ x+2$ *e si procede con* $w_0=2+¼\ 2\ (4-4)=2$, $w_1= 2+¼\ 2\ (4-4)\ 1=2$.
d) $h(x)= 2x + 2$. Disegnarla. Si può osservare che passa anch'essa per (1,4).

2- Le borsette di Maria 2018

Maria deve scegliere una nuova borsetta. Sapete che gradisce una borsa chiara piccola ma non gradisce né una borsetta scura piccola, né una chiara grande. Pensate che comprerebbe una borsetta grande scura?
Proporre una soluzione in base a diversi modelli, codificando il colore chiaro con 0 e quello scuro con 1 e la dimensione piccola con 0 e grande con 1. In particolare:

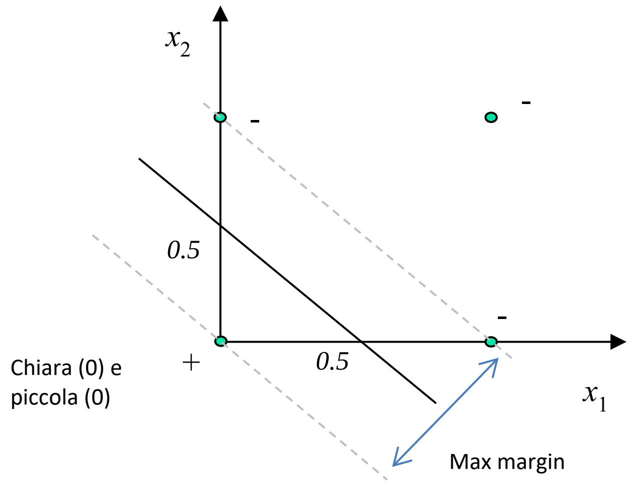
1. Definire il compito (task) di apprendimento (input, output, data set e tipo di task)
2. Proporre la soluzione che forniremmo costruendo l'ipotesi con Find-S (mostrando i passi di esecuzione dell'algoritmo per costruire l'ipotesi)
3. Proporre una soluzione utilizzando Candidate Elimination
4. Proporre una soluzione mostrando la costruzione con ID3 tramite il calcolo dei Gain
5. Proporre una soluzione (direttamente, senza necessità di eseguire algoritmi di learning) con un modello lineare (per via grafica ed esprimendo la $h(\mathbf{x})$)
6. Proporre una soluzione (direttamente, senza necessità di eseguire algoritmi di learning) con un modello SVM *hard-margin* (per via grafica ed esprimendo la $h(\mathbf{x})$, motivandola)
7. Formulare la risposta che si darebbe con un K-NN (motivandola)
8. Formulare la risposta che si darebbe con un modello di tipo "Lookup table" (motivandola)

Soluzioni (solo risultato per verifica, svolgere personalmente)

1. $\mathbf{x}= [x_1, x_2]$, $x_i \in \{0,1\}$ $i=1,2$. Output e target $y_p \in \{0,1\}$. Data set $(\mathbf{x}_p, y_p)$ $p=1..l$ (con l=3 in questo task). Task di classificazione binaria.
2. Costruire prima la tabella di verità. Ovviamente senza avere una risposta in [1,1].
   Si ottiene h=<0,0> (i passi dell'algoritmo sui 3 esempi dati sono lasciati da scrivere per esercizio).
   Risponde quindi 0, ossia negativamente in [1,1] (non compra la borsetta)
3. S=G=<0,0>. Risponde quindi negativamente in [1,1] (non compra la borsetta)
4. Si ottiene dopo il calcolo: G(S, x_1) = E(S)-2/3. G(S, x_2) = E(S)-2/3. Sono pari: si sceglie a caso. Disegno alberi lasciato per esercizio. Si ottiene: Un albero radicato in x_1 con ramo zero che interroga x_2 avendo poi 1 solo se anch'essa è zero. Oppure un albero radicato in x_2 che interroga x_1 avendo poi 1 solo se anch'essa è zero. Risponde 0, negativamente, in [1,1] (non compra la borsetta).
5. Decision boundary: $- x_1 - x_2 + 0.5=0$. $h(\mathbf{x})$= 1 se $(- x_1 - x_2 + 0.5 \geq 0)$, 0 altrimenti. Disegno lasciato per esercizio (pag. successiva). In [1,1] (applicarlo a $h(\mathbf{x})$) risponde 0 (o negativamente se si usa *sign()*), (ossia non compra la borsetta). Nota: è il NOR visto alla prima esercitazione!!!!!
6. La stessa risposta fornita al punto 5, osservando che nel disegno il decision boundary (tra quelli possibili) è quello che massimizza il margine in questo task (passa in mezzo tra i punti vicini al decision boundary disegnati nel piano 2D al punto precedente per le zone positive e negative). Vedere figura a pagina successiva. Risponde 0, o negativamente, in [1,1] (non compra la borsetta).
7. Si ha risposta negativa (non compra la borsetta) perché per l'equazione del K-NN (scriverla) per 1-NN, 2-NN, e 3-NN si ha sempre una maggioranza di punti zero (o negativi) nell'intorno del punto [1,1] (scrivere i 3 casi con 1 -; 2 -; 2 - e 1 +, rispettivamente nell'intorno, ossia K=1, 2 o 3) .
8. Nessuna risposta, poiché non trova [1,1] in tabella.

P.S. In molti avevano risposto scrivendo le $h(\mathbf{x})$ ma dimenticato di rispondere anche alla domanda: "Pensate che comprerebbe una borsetta grande scura?". Quindi è suggerito di (ri)leggere bene il testo.

Si noti anche che non tutti i punti sono in sequenza per cui ad alcuni si può rispondere correttamente anche se non fosse riuscita la soluzione di alcuni punti precedenti (e.g. il 7 e 8 rispetto a 2, 3, 4, 5, 6).


x_2
-
-
0.5
-
Chiara (0) e piccola (0)
+
0.5
x_1
Max margin