

Εισαγάγετε την απάντησή σας

4

Ιδιότητα *

ΔΠΜΣ, Υ.Δ. κοκ

Εισαγάγετε την απάντησή σας

5

$$\begin{aligned} i_t &= \sigma(W^{(i)}x_t + U^{(i)}h_{t-1}) \\ f_t &= \sigma(W^{(f)}x_t + U^{(f)}h_{t-1}) \\ o_t &= \sigma(W^{(o)}x_t + U^{(o)}h_{t-1}) \\ \tilde{c}_t &= \tanh(W^{(c)}x_t + U^{(c)}h_{t-1}) \\ c_t &= f_t \circ c_{t-1} + i_t \circ \tilde{c}_t \\ h_t &= o_t \circ \tanh(c_t) \end{aligned}$$

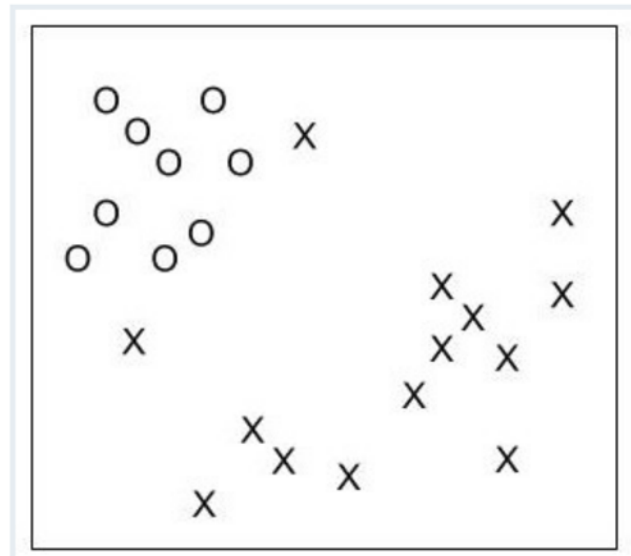
Οι εξισώσεις ενός κυττάρου (cell) LSTM είναι αυτές τις εικόνες.

Γράψτε σωστό ή λάθος για τις επόμενες προτάσεις με συντομοτάτη αιτιολόγηση.

- α) αν $x_t = 0$ τότε $h_t = h_{t-1}$
 - β) αν το f_t είναι πολύ μικρό ή μηδέν τότε το σφάλμα δεν θα διαδοθεί σε προηγούμενα χρονικά βήματα
 - γ) f_t, i_t, o_t είναι μη αρνητικές
 - δ) μπορούμε να θεωρήσουμε τις f_t, i_t, o_t ως κατανομές πιθανότητας (μη αρνητικές και να αθροίζουν στο 1)
- (3 βαθμοί)

Εισαγάγετε την απάντησή σας

6



Δίνεται το binary dataset του σχήματος. Διαθέτουμε ένα SVM με τετραγωνικό πυρήνα δηλαδή δευτεροβάθμιο πολυωνυμικό πυρήνα. Η παράμετρος κόστους " C " καθορίζει τη θέση και τη μορφή του διαχωριστικού συνόρου μεταξύ των δύο κλάσεων.

Που θα είναι το διαχωριστικό σύνορο για τιμές του C που τείνουν στο 0;
Απαντήστε σε μια πρόταση και σχεδιάστε το πάνω στο σχήμα (ξανασχεδιάστε το πρόχειρα). (Μη ανώνυμη ερώτηση ⓘ)
(3 βαθμοί)

↑ Αποστολή αρχείου

Όριο πλήθους αρχείων: 1 Όριο μεγέθους μεμονωμένου αρχείου: 10MB Επιτρεπόμενοι τύποι αρχείων: Word, Excel, PPT, PDF, Εικόνα, Βίντεο, Ήχος

7

Έστω ότι έχουμε πρόβλημα δυαδικής ταξινόμησης σε χώρο δύο διαστάσεων με τα δείγματα $(-1,0)$, $(-1,9)$, $(10,0)$, $(10,10)$ και (a,b) να ανήκουν στην κλάση -1 και τα δείγματα $(4.5,4)$, $(4.5,5)$, $(5.5,4)$, $(5.5, 5)$ και (c,d) ανήκουν στην κλάση $+1$. Να βρείτε τον βέλτιστο PDS πυρήνα που καθιστά τις κλάσεις διαχωρίσιμες και να τεκμηριώσετε την επιλογή σας. Τα $abcd$ είναι τα 4 τελευταία ψηφία του αριθμού μητρώου σας.
(7 βαθμοί)

Εισαγάγετε την απάντησή σας

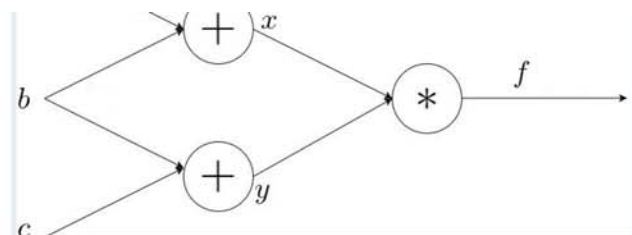
8

Ποια από τις ακόλουθες προτάσεις δεν είναι αληθής για έναν εκπαιδευμένο αυτοοργανούμενο χάρτη (SOM);
(3 βαθμοί)

- ☐ Ο χάρτης πραγματοποιεί διανυσματικό κβαντισμό του χώρου εισόδου.
- ☐ Δεν απαντώ
- ☐ Κατά την απεικόνιση της εισόδου στην έξοδο, η μείωση της διάστασης εξασφαλίζει τη διατήρηση της διάταξης.
- ☐ Η απεικόνιση της εισόδου στην έξοδο μπορεί να παρασταθεί με τη βοήθεια ενός διαγράμματος Voronoi.
- ☐ Καμία από τις υπόλοιπες απαντήσεις (είναι όλες αληθείς).
- ☐ Συνήθως, η απεικόνιση της εισόδου στην έξοδο χαρακτηρίζεται από μείωση της διάστασης.

9





Δίνεται ο υπολογιστικός γράφος του σχήματος. Αν $a=3$, $b=-2$, $c=2$ υπολογίστε τις μερικές παραγώγους προς όλες τις μεταβλητές του γράφου χρησιμοποιώντας τον κανόνα της αλυσίδας. (Μη ανώνυμη ερώτηση ①)
(5 βαθμοί)

↑ Αποστολή αρχείου

Όριο πλήθους αρχείων: 1 Όριο μεγέθους μεμονωμένου αρχείου: 10MB Επιτρεπόμενοι τύποι αρχείων: Word, Excel, PPT, PDF, Εικόνα, Βίντεο, Ήχος

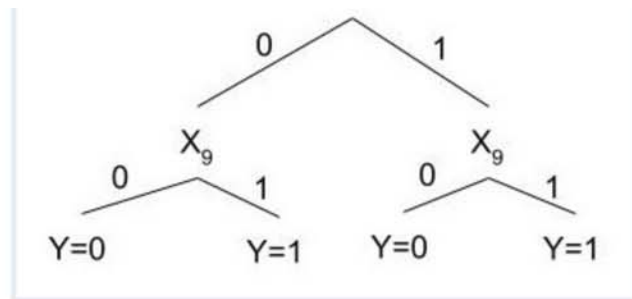
10

Έστω ότι έχουμε ένα σύστημα online μάθησης το οποίο καλείται να επιλέξει από 5 πιθανές δράσεις (με κωδικό 0 ως 4). Να γράψετε την κατανομή p_2 πάνω στις δράσεις καθώς και τα βάρη $w_{2,i}$ των δράσεων που προκύπτουν μετά την εφαρμογή του πρώτου γύρου του randomized αλγορίθμου σταθμισμένης πλειοψηφίας, αν το διάνυσμα απώλειας l_1 του πρώτου γύρου έχει μη-μηδενική τιμή για τις ενέργειες $(100 + A) \bmod 5$ και $(A - 99) \bmod 5$, όπου A είναι τα 2 τελευταία ψηφία του αριθμού μητρώου σας. Θεωρείστε επίσης ότι η παράμετρος β είναι ίση με 0.9, όπου $x=9-k$ και k το τελευταίο ψηφίο του αριθμού μητρώου σας (θέστε $x=0.6$ αν το τελευταίο ψηφίο του αριθμού μητρώου σας είναι ίσο με το 9).
(5 βαθμοί)

Εισαγάγετε την απάντησή σας

11

X_5



Θεωρούμε έναν αλγόριθμο μάθησης βασισμένο σε δέντρα αποφάσεων (ΔΑ) ο οποίος στην είσοδο δέχεται δυαδικά χαρακτηριστικά X_i και στην έξοδο παράγει μια επίσης δυαδική απόφαση ($Y \in \{0,1\}$). Ο συγκεκριμένος αλγόριθμος παράγει μόνο κανονικά (regular) ΔΑ με βάθος 2. Ένα κανονικό ΔΑ βάθους δύο είναι ένα ΔΑ με βάθος δύο (δηλαδή τέσσερα φύλλα) στο οποίο το αριστερό και το δεξί παιδί της ρίζας υποχρεωτικά ελέγχουν το ίδιο χαρακτηριστικό. Στο σχήμα δίνεται ένα παράδειγμα τέτοιου ΔΑ.

α) Θεωρήστε ότι έχετε δεδομένα για μια έννοια-στόχο C η οποία μπορεί να περιγραφεί με κανονικά ΔΑ βάθους δύο. Θεωρήστε επιπρόσθετα ότι έχετε 10 χαρακτηριστικά εισόδου (από τα οποία τελικά θα χρειαστούν μόνο δύο για την κατασκευή του ΔΑ). Πόσα δείγματα θα χρειαστούν στην εκπαίδευση έτσι ώστε με πιθανότητα 98% ο αλγόριθμος να βρίσκει ένα δέντρο με ορθότητα το λιγότερο 97%; (3 μονάδες)

β) Ξεκινώντας από την παρατήρηση ότι πολλές υποθέσεις είναι ισοδύναμες, δηλαδή κάνουν την ίδια αντιστοίχιση από τον χώρο των χαρακτηριστικών στον χώρο εξόδου -για παράδειγμα η σειρά των δύο χαρακτηριστικών δεν έχει σημασία-, βρείτε ένα πιο σφιχτό φράγμα που να απαιτεί λιγότερα δεδομένα για τις ίδιες τιμές πιθανότητας και ορθότητας.

(4 μονάδες) (Μη ανώνυμη ερώτηση ⓘ)

(7 βαθμοί)

↑ Αποστολή αρχείου

Όριο πλήθους αρχείων: 1 Όριο μεγέθους μεμονωμένου αρχείου: 10MB Επιτρεπόμενοι τύποι αρχείων: Word, Excel, PPT, PDF, Εικόνα, Βίντεο, Ήχος

12

Έστω ότι έχουμε πρόβλημα δυαδικής ταξινόμησης (ετικέτες $\{-1, +1\}$), το οποίο επιθυμούμε να μάθουμε τη χρήση ασθενών μοντέλων μάθησης και της τεχνικής AdaBoost. Έστω επίσης ότι διαθέτουμε ένα σύνολο 5 δειγμάτων $S = \{(x_0, y_0) \dots (x_4, y_4)\}$ του εν λόγω προβλήματος και ότι μετά την πρώτη επανάληψη του αλγορίθμου το σφάλμα ταξινόμησης είναι ίσο με 0.χ, όπου χ το τελευταίο ψηφίο του αριθμού μητρώου σας (θεωρείστε $x=3$ στην περίπτωση που το τελευταίο ψηφίο του αριθμού μητρώου σας είναι το 0). Έστω επίσης ότι τα δείγματα $(100 + A) \bmod 5$ και $(A - 99) \bmod 5$ ταξινομούνται εσφαλμένα από το ασθενές μοντέλο μάθησης h_1 , ενώ τα υπόλοιπα ταξινομούνται ορθά (Α είναι τα 2 τελευταία ψηφία του αριθμού μητρώου σας). Πως θα διαμορφωθεί η κατανομή επιλογής των δειγμάτων (D2) στη δεύτερη επανάληψη του AdaBoost;
(5 βαθμοί)

Εισαγάγετε την απάντησή σας

13

Μια συνάρτηση h που ορίζεται όπως φαίνεται παρακάτω είναι συμμετρική αν η τιμή της εξαρτάται μόνο από το πλήθος των 1 στην είσοδο. Έστω S το σύνολο όλων των συμμετρικών συναρτήσεων. Ποια είναι η διάσταση VC του S ;
(5 βαθμοί)

$$h : \{0, 1\}^n \rightarrow \{0, 1\}$$

Εισαγάγετε την απάντησή σας

14

Η συνάρτηση ενεργοποίησης ReLU (z) = $\max(0, z)$ μπορεί να σβήσει/κορεστεί όταν η είσοδος είναι αρνητική. Σας προτείνεται η χρήση μιας άλλης συνάρτησης ενεργοποίησης $f(z) = 1.5z$. Θα λυθεί έτσι το πρόβλημα? Εξηγείστε.
(3 βαθμοί)

Εισαγάγετε την απάντησή σας

15

$$h_r(x) = 1 \text{ if } x \leq r \text{ και } h_r(x) = 0 \text{ αλλιώς (1)}$$

Έχουμε μονοδιάστατα δεδομένα. Για έναν πραγματικό r , ορίζουμε τη συνάρτηση (1). Έστω $H = \{h_r\}$. Θεωρούμε ένα σύνολο S από m ξεχωριστά δείγματα πάνω στην ευθεία. Ποια είναι η εμπειρική πολυπλοκότητα Rademacher $R_m(H)$ της H στο S ; (Μη ανώνυμη ερώτηση ⓘ)
(5 βαθμοί)

↑ Αποστολή αρχείου

Όριο πλήθους αρχείων: 1 Όριο μεγέθους μεμονωμένου αρχείου: 10MB Επιτρεπόμενοι τύποι αρχείων: Word, Excel, PPT, PDF, Εικόνα, Βίντεο, Ήχος

16

Εξηγήστε εν συντομία σε τί διαφέρουν το bipartite ranking από το k-partite ranking και δώστε ένα παράδειγμα χρήσης (ένα υποθετικό task) για το καθένα.
(3 βαθμοί)

Εισαγάγετε την απάντησή σας

17

Ας υποθέσουμε ότι εκπαιδεύετε ένα νευρωνικό δίκτυο για ταξινόμηση, αλλά παρατηρείτε ότι το training error είναι πολύ χαμηλότερο από το validation error. Ποιο από τα ακόλουθα θα χρησιμοποιήσετε για την αντιμετώπιση του ζητήματος (επιλέξτε όλα όσα ισχύουν);
(3 βαθμοί)

- ☐ Μειώστε την dropout probability
- ☐ Χρησιμοποιήστε ένα δίκτυο με λιγότερα επίπεδα
- ☐ Αυξήστε το βάρος της L2 κανονικοποίησης (regularization)
- ☐ Αυξήστε το μέγεθος κάθε κρυμμένου επιπέδου

18

Στη γραμμική παλινδρόμηση, έχουμε $XX^T W = XY$. Τί πρέπει να ισχύει για να είναι αντιστρέψιμος ο πίνακας XX^T ;
(3 βαθμοί)

Εισαγάγετε την απάντησή σας

19

Έχουμε μια συλλογή εικόνων 32×32 pixels σε χρωματική αναπαράσταση RGB. Για την ταξινόμησή τους σε 5 κατηγορίες σας δίνεται συνελκτικό δίκτυο με τα εξής επίπεδα: CONV (8 φίλτρα, $F=6$, $S=2$, $P=1$), POOL ($F=3$, $S=2$) FC (32), Softmax(5). Υπολογίστε τον αριθμό των παραμέτρων του δικτύου. (5 βαθμοί)

Εισαγάγετε την απάντησή σας

20

Έστω ότι για ένα πρόβλημα δυαδικής ταξινόμησης έχουμε 100 δείγματα δεδομένων με τις αντίστοιχες ετικέτες τους. Το προσεγγίζουμε με ένα πολυωνμικό μοντέλο 20ου βαθμού και παρατηρούμε ότι εμφανίζει υπερπροσαρμογή. Για το λόγο αυτό αποφασίζουμε να εφαρμόσουμε L2 κανονικοποίηση με $\lambda = 0.xy$, όπου x, y το προτελευταίο και το τελευταίο ψηφίο του αριθμού μητρώου σας αντίστοιχα (θέστε $\lambda = 0.27$ αν $xy = 00$). Πως θα μεταβάλλονται οι οροι του πολυωνύμου σε κάθε επανάληψη, αν χρησιμοποιήσουμε ως μέθοδο εκμάθησης τη στοχαστική κατάβαση κλήσης με ρυθμό μάθησης $\eta = 0.y$ (θέστε $\eta = 0.7$ αν $y = 0$); (3 βαθμοί)

Εισαγάγετε την απάντησή σας

21

Ένα αυτόνομο ρομπότ κινείται στο χώρο που απεικονίζεται στο σχήμα (α). Οι καταστάσεις αντιστοιχούν στα κελιά και αναφερόμαστε σε αυτά με τη σύμβαση (γραμμή, κολόνα). Το ρομπότ ξεκινάει πάντα από την κατάσταση "S". Υπάρχουν δύο τερματικές καταστάσεις-στόχοι στις θέσεις (1,3) και (2,3) στις οποίες αναγράφεται η ανταμοιβή του πράκτορα. Στις μη-τερματικές θέσεις η ανταμοιβή είναι μηδέν. Ο πράκτορας μπορεί να κινηθεί σε τέσσερις κατευθύνσεις (πάνω, κάτω, αριστερά, δεξιά). Ωστόσο, η κίνησή του είναι ελαφρά στοχαστική: μετακινείται με πιθανότητα 0.8 στην κατεύθυνση που θέλει, αλλά και με πιθανότητα 0.1 για την καθεμία, προς τις δύο κατευθύνσεις που είναι κάθετες στην ηθελημένη (σχήμα β). Αν το ρομπότ συγκρουστεί με τοίχο (πάει να βγει εκτός του grid), μένει στην ίδια θέση.

1) αν ο παράγοντας έκπτωσης είναι 0.9 υπολογίστε τις τιμές της αξίας για τα σημεία (2,1), (1,2) και (2,2) για δύο επαναλήψεις του αλγόριθμου value iteration (4 μονάδες)

2) ο πράκτορας ξεκινάει με την πολιτική που διαλέγει να πηγαίνει πάντοτε δεξιά. Εκτελεί τρία πειράματα και οι μετακινήσεις του καταγράφονται ως εξής: α) (1,1)-(1,2)-(1,3), β) (1,1)-(1,2)-(2,2)-(2,3), και γ) (1,1)-(2,1)-(2,2)-(2,3). Ποια είναι η κατά monte carlo εκτίμηση της αξίας των καταστάσεων (1,1) και (2,2) με βάση αυτά τα δειγματικά μονοπάτια; (3 μονάδες)

(Μη ανώνυμη ερώτηση ①)

(7 βαθμοί)