ΑΝΑΓΝΩΡΙΣΗ ΠΡΟΤΥΠΩΝ

Τμήμα Ηλεκτρολόγων Μηχανικών και Τεχνολογίας Υπολογιστών

Εξεταστική περίοδος Ιουνίου 1997

1 Μονάδες 3

Δίνονται τα ακόλουθα παραδείγματα δύο κατηγοριών:

Κατηγορία	π_1	π_2	π_3	π_4
ω_1	-3,-1	1,1	2,0	3,1
ω_2	-2,0	0,0	1,-1	

α. Υπολογίστε τους συντελεστές βαρύτητας της γραμμικής συνάρτησης απόφασης χρησιμοποιώντας την μέθοδο των Kiefer-Wolfowitz. Εκτελέστε μόνο δύο βήματα του επαναληπτικού τμήματος του αλγόριθμου.

β. Ποιό είναι το ελάχιστο σφάλμα του συστήματος ταξινόμησης προτύπων που μπορεί να επιτύχει η μέθοδος ταξινόμησης με γραμμικές συναρτήσεις απόφασης;

γ. Πώς μπορείτε να βελτιώσετε την αξιοπιστία του συστήματος ταξινόμησης;

Λύση του α:

Επεκτείνω την διάσταση των προτύπων $\mathbf{x} = (\mathbf{x}, 1)^T$, οπότε και τα πρότυπα των παραδειγμάτων γίνονται:

$$\Omega_1 = \{(-3,-1,1),(1,1,1),(2,0,1),(3,1,1)\}$$

$$\Omega_2 = \{(-2,0,1),(0,0,1),(1,-1,1)\}$$

Η συνάρτηση απόφασης δίνεται από την σχέση:

$$g(\mathbf{x}) = p(\omega|\mathbf{x}) \approx \sum_{i=1}^{Q} c_i \varphi_i(\mathbf{x}) = \mathbf{c}^T \mathbf{x}_\varphi$$

Εφόσον δεν πραγματοποιήσουμε κάποιον μη-γραμμικό μετασχηματισμό, η συνάρτηση απόφασης δίνεται από την σχέση:

$$g(\mathbf{x}) = p(\omega|\mathbf{x}) \approx \mathbf{c}^T\mathbf{x}$$

Φτιάχνω την συνάρτηση που δίνει την διαφορά των συναρτήσεων απόφασης:

$$d(\mathbf{x}) = g_1(\mathbf{x}) - g_2(\mathbf{x}) = p(\omega_1|\mathbf{x}) - p(\omega_2|\mathbf{x}) \approx \mathbf{c_1}^T\mathbf{x} - \mathbf{c_2}^T\mathbf{x} = (\mathbf{c_1}^T - \mathbf{c_2}^T)\mathbf{x} = \mathbf{c}^T\mathbf{x}$$

Αντί λοιπόν για έξι συντελεστές που πρέπει να υπολογίσουμε, αν ορίσουμε δύο συναρτήσεις απόφασης, μπορούμε να χρησιμοποιήσουμε την συνάρτηση $d(\mathbf{x})$ και να ορίσουμε την εξής παλίνδρομη συνάρτηση:

$$z(\mathbf{x}) = \begin{cases} 1 & \mathbf{x} \in \omega_1 \\ -1 & \mathbf{x} \in \omega_2 \end{cases}$$

Ο στοχαστικός αλγόριθμος επαναπροσδιορισμού των συντελεστών βαρύτητας γίνεται:

$$\mathbf{c}_{k+1} = \begin{cases} \mathbf{c}_k + a(k)\mathbf{x}, & z(\mathbf{x}) > \mathbf{c}^T\mathbf{x} \\ \mathbf{c}_k - a(k)\mathbf{x}, & z(\mathbf{x}) \leq \mathbf{c}^T\mathbf{x} \end{cases}$$

Θα εφαρμόσουμε τον αλγόριθμο των Kiefer-Wolfowitz χρησιμοποιώντας τις ακολουθίες πραγματικών αριθμών

$$a(k) = \frac{1}{k}, \quad k = 1, 2, ...$$

Ξεκινώ με αρχική τιμή των συντελεστών $c_1 = (0,0,0)^T$.
Γιά το πρώτο παράδειγμα, το $(-3,-1,1)$, έχω:

$$g((-3,-1,1) = (0,0,0)(-3,-1,1)^T = 0 \quad < \quad z((-3,-1,1)) = 1$$

Οι νέες τιμές των συντελεστών βαρύτητας γίνονται:

$$\mathbf{c}_2 = \mathbf{c}_1 + a(1)(-3,-1,1)^T = (-3,-1,1)^T$$

Γιά το δεύτερο παράδειγμα, το $(1,1,1)$, έχω:

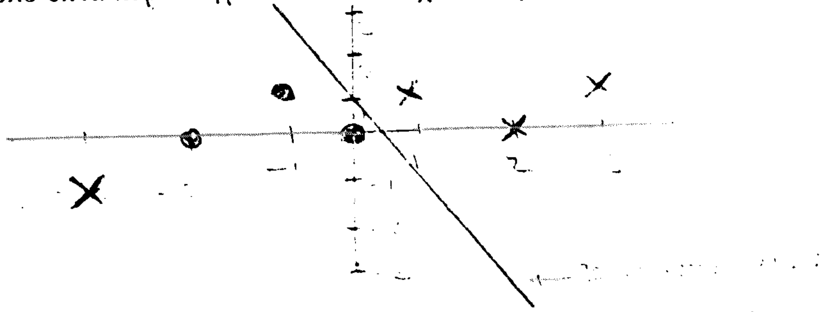
$$g((1,1,1) = (-3,-1,1)(1,1,1)^T = -3 \quad < \quad z((1,1,1)) = 1$$

Οι νέες τιμές των συντελεστών βαρύτητας γίνονται:

$$\mathbf{c}_3 = \mathbf{c}_2 + a(2)(1,1,1)^T = (-2.5,-0.5,1.5)^T$$

Λύση του β:

Αν απεικονίσουμε τα παραδείγματα στο επίπεδο θα δούμε ότι τα πρότυπα των δύο κατηγοριών δεν είναι γραμμικά διαχωρίσιμα. Η ευθεία που διαχωρίζει τα πρότυπα των παραδειγμάτων κατά τον καλύτερο δυνατό τρόπο ταξινομεί το πρότυπο (-3,-1) στην δεύτερη κατηγορία. Συνεπώς έχουμε ένα σφάλμα σε ένα σύνολο επτά παραδειγμάτων. Το ελάχιστο σφάλμα είναι $1/7 = 14.28\%$.



Λύση του γ:

Βελτίωση της αξιοπιστίας του συστήματος ταξινόμησης μπορούμε να πετύχουμε με την μέθοδο της μη-γραμμικής επέκτασης του διανύσματος των προτύπων.

2 Μονάδες 3

Διπλή επιδημία των θανατηφόρων ασθενειών χολέρας και ηπατίτιδας κτύπησαν ένα αποκλεισμένο νησί 3000 κατοίκων στο οποίο δεν υπάρχει γιατρός, τηλεπικοινωνικά μέσα και τυχαίνει να είστε ο κυβερνήτης του νησιού.

Το μοναδικό ιατρικό διαγνωστικό όργανο που διαθέτετε είναι ένα θερμόμετρο. Από ιατρικά βιβλία συγκεντρώσατε τις ακόλουθες πληροφορίες:

α. Ενας υγιής άνθρωπος έχει θερμοκρασία σώματος στο διάστημα θερμοκρασιών 36.3-37 βαθμούς Κελσίου.

β. Μετρήσεις που έγιναν σε ανθρώπους που έπασχαν από χολέρα έδωσαν μέση τιμή θερμοκρασίας σώματος 38 βαθμούς Κελσίου με διασπορά 1.

γ. Μετρήσεις που έγιναν σε ανθρώπους που έπασχαν από ηπατίτιδα έδωσαν μέση τιμή θερμοκρασίας σώματος 39 βαθμούς Κελσίου με διασπορά 1.5.

Εχετε στην διάθεσή σας ένα φάρμαχο γιά κάθε ασθένεια. Γνωρίζετε από την βιβλιογραφία ότι μόνο ένα από τα δύο φάρμακα μπορείτε να δώσετε σε ασθενή. Η ταυτόχρονη χορήγηση τους είναι θανατηφόρα. Γνωρίσετε επίσης ότι η ύπαρξη της μιάς ασθένειας αποκλείει την ύπαρξη της άλλης. Το φάρμακο γιά την χολέρα έχει επιτυχία στο 75% των περιπτώσεων ενώ το φάρμακο γιά την ηπατίτιδα γιατρεύει το 98% των ασθενών. Αν δώσετε ένα από τα φάρμακα σε υγιή άνθρωπο τότε αυτή η πράξη δεν έχει καμμιά επίδραση στην υγεία του.

Μπορείτε να βρείτε μία μέθοδο με την οποία να αποφασίζετε την χορήγηση κάποιου φαρμάκου στον πληθυσμό του νησιού βάσει της σωματικής του θερμοκρασίας έτσι ώστε να ελαχιστοποιήσετε τον αριθμό των θανάτων. Χρησιμοποιείστε την υπόθεση ότι το ένα τρίτο των κατοίκων έχει προσβληθεί από το μικρόβιο της ηπατίτιδας και το ένα δέκατο από το μικρόβιο της χολέρας. Κάποια από τα δεδομένα που σας δίνονται ενδεχομένως να μην είναι απαραίτητα γιά την λύση του προβλήματός σας.

Λύση:

Εφόσον η χορήγηση των φαρμάκων σε υγιή άνθρωπο δεν έχει καμμία επίπτωση στην υγεία του τότε θα χορηγήσουμε οποσδήποτε ένα από τα δύο φάρμακα σε κάθε κάτοικο. Το πρόβλημα συνεπώς είναι να αποφασίσουμε από ποιά ασθένεια είναι πιό πιθανό να πάσχει ο άνθρωπος από τον οποίο θα πάρουμε την θερμοχρασία του σώματος.

Κριτήριο απόφασης είναι η μεγιστοποίηση του αριθμού ασθενών που θα γιατρευτούν. Ο αριθμός αυτός μπορεί να αναλυθεί στους εξής δύο όρους:

$$P(Ανάρρωση\ ασθενή) = P(Απόφαση\ χολέρα, Πάσχει\ από\ χολέρα, Επιτυχία\ φαρμάκου\ χολέρας)$$

$$+P(Απόφαση\ ηπατίτιδα, Πάσχει\ από\ ηπατίτιδα, Επιτυχία\ φαρμάκου\ ηπατίτιδας) =$$

Επειδή η πιθανότητα επιτυχούς ανάρρωσης είναι σταθερή (δίνεται από τα δεδομένα της άσκησης) και δεν εξαρτάται από άλλους παράγοντες, η πιθανότητα επιτυχούς ανάρρωσης αναλύεται στους ακόλουθους όρους:

$$P(Ανάρρωση\ ασθενή) = P(Απόφαση\ ηπατίτιδα | Πάσχει\ από\ ηπατίτιδα)$$

$$P(Πάσχει\ από\ ηπατίτιδα)P(Επιτυχία\ φαρμάκου\ ηπατίτιδας)$$

$$+P(Απόφαση\ χολέρα | Πάσχει\ από\ χολέρα)P(Πάσχει\ από\ χολέρα)P(Επιτυχία\ φαρμάκου\ χολέρας)$$

Αντικαθιστώντας τις γνωστές τιμές έχω:

$$P(Ανάρρωση\ ασθενή) = P(Απόφαση\ ηπατίτιδα | Πάσχει\ από\ ηπατίτιδα) 0.3333 * 0.98$$

$$+P(Απόφαση\ χολέρα | Πάσχει\ από\ χολέρα) * 0.1 * 0.75$$

Οι δεσμευμένες πιθανότητες υπολογίζονται από τις αντίστοιχες πυκνότητες πιθανότητας. Με κριτήριο την μεγιστοποίηση της πληροφορίας που λαμβάνουμε από τα δεδομένα του έχουμε στην διάθεσή μας οι αντίστοιχες πυκνότητες πιθανότητας υποθέτουμε ότι ακολουθούν την κανονική κατανομή (θεώρημα 6, σελίδα 77).

Υποθέτοντας ότι μετρήσαμε την θερμοχρασία ενός ανθρώπου και την βρήκαμε να είναι Θ. Ψάχνοντας γιά το μέγιστο της συνάρτησης $P(Ανάρρωση\ ασθενή)$ έχουμε:

$$\frac{\partial P(Ανάρρωση\ ασθενή)}{\partial \Theta} = 0$$

$$\frac{\partial}{\partial \Theta} \frac{0.3333 * 0.98}{\sqrt{2\pi}\sqrt{1.5}} \int_{-\infty}^{\Theta} e^{-\frac{(x-39)^2}{2\,1.5^2}} + \frac{\partial}{\partial \Theta} \frac{0.1 * 0.75}{\sqrt{2\pi}} \int_{\Theta}^{+\infty} e^{-\frac{(x-38)^2}{2}} = 0$$

Απλοποιώντας τελικά οδηγούμαστε στην επίλυση της δευτεροβάθμιας εξίσωσης:

$$\Theta^2 - 72\Theta + 1282.38774 = 0$$

Εχουμε δύο λύσεις, τις

$$\Theta_1 = 39.68 \;\; \kappa\alpha\iota \;\; \Theta_2 = 32.31$$

Συνεπώς όταν μετρήσουμε θερμοκρασία μικρότερη των 39.68 και μεγαλύτερη των 32.31 βαθμών κελσίου χορηγούμε το φάρμακο γιά την χολέρα. Διαφορετικά χορηγούμε το φάρμακο γιά την ηπατίτιδα. Πρακτικά το κάτω όριο των 32.31 βαθμών δεν το λαμβάνουμε υπόψη.

Με το κριτήριο ταξινόμησης που βρήκαμε θα πετύχουμε να γιατρέψουμε τον μεγαλύτερο αριθμό ανθρώπων με τα δεδομένα που διαθέτουμε και τις υποθέσεις που κάναμε.

3 Μονάδες 2

Δίνονται τα ακόλουθα παραδείγματα δύο κατηγοριών:

Κατηγορία	π_1	π_2	π_3	π_4
ω_1	-3,-1	1,1	2,0	3,1
ω_2	-2,0	0,0	1,-1	

α. Κατασκευάστε σύστημα ταξινόμησης προτύπων με την βοήθεια συναρτήσεων δυναμικού. Εκτελέστε όλα τα βήματα του αλγόριθμου μέχρι να χρησιμοποιήσετε μία φορά όλα τα παραδείγματα που έχετε στην διάθεσή σας.

β. Ποιά είναι τα πλεονεκτήματα του συστήματος ταξινόμησης προτύπων με συναρτήσεις δυναμικού σε σχέση με σύστημα ταξινόμησης που χρησιμοποιεί γραμμικές συναρτήσεις απόφασης. Ποιό είναι το πλέον κατάλληλο σύστημα γιά την περίπτωση της άσκησης; (Αιτιολογήστε την απάντησή σας).

Λύση α:

Επειδή έχω την ελευθερία επιλέγω την υπολογιστικά απλούστερη συνάρτηση δυναμικού:

$$K(\mathbf{x}, \mathbf{x}_k) = \frac{1}{1 + |\mathbf{x} - \mathbf{x}_k|}$$

Αρχικά θέτω $g_{12}(\mathbf{x}) = 0$ και ορίζω να είναι $g_{12}(\mathbf{x}) = -g_{21}(\mathbf{x})$.
Υπολογίζω την τιμή της συνάρτησης απόφασης γιά το πρώτο παράδειγμα, το (-3,-1):

$$g_{12}((-3,-1)) = 0 \;\; \Rightarrow \;\; r_k = 1$$

$$g_{12}(\mathbf{x}) = \frac{1}{1 + |\mathbf{x} - (-3,-1)|}$$

Ολα τα παραδείγματα θα πάρουν θετική τιμή με την τρέχουσα συνάρτηση απόφασης. Συνεπώς δεν χρειάζεται επαναπροσδιορισμός της συνάρτησης απόφασης γιά κανένα από τα παραδείγματα της πρώτης κατηγορίας

Υπολογίζω την τιμή της συνάρτησης απόφασης γιά το πρώτο παράδειγμα της δεύτερης κατηγορίας, το (-2,0)

$$g_{12}((-2,0)) > 0 \quad \Rightarrow \quad r_k = -1$$

$$g_{12}(\mathbf{x}) = \frac{1}{1+|\mathbf{x}-(-3,-1)|} - \frac{1}{1+|\mathbf{x}-(-2,0)|}$$

Υπολογίζω την τιμή της συνάρτησης απόφασης γιά το δεύτερο παράδειγμα της δεύτερης κατηγορίας, το (0,0)

$$g_{12}((0,0)) = \frac{1}{1+\sqrt{10}} - \frac{1}{1+2} < 0 \quad \Rightarrow \quad r_k = 0$$

Δεν χρειάζεται επαναπροσδιορισμός.

Υπολογίζω την τιμή της συνάρτησης απόφασης γιά το τρίτο παράδειγμα της δεύτερης κατηγορίας, το (1,-1)

$$g_{12}((1,-1)) = \frac{1}{1+4} - \frac{1}{1+\sqrt{2}} < 0 \quad \Rightarrow \quad r_k = 0$$

Δεν χρειάζεται επαναπροσδιορισμός. Η τελική συνάρτηση απόφασης δίνεται από την σχέση:

$$g_{12}(\mathbf{x}) = \frac{1}{1+|\mathbf{x}-(-3,-1)|} - \frac{1}{1+|\mathbf{x}-(-2,0)|}$$

Λύση β:

Τα συστήματα ταξινόμησης προτύπων με συναρτήσεις δυναμικού είναι μη-γραμμικά συστήματα. Γιαυτό τον λόγο έχουν την δυνατότητα να ταξινομούν σωστά και πρότυπα κατηγοριών τα οποία δεν είναι γραμμικά διχωρίσιμα. Οριακά τα συστήματα αυτά μπορούν να ελαχιστοποιήσουν το σφάλμα ταξινόμησης.

Στο συγκεκριμένο παράδειγμα επειδή τα παραδείγματα δεν είναι γραμμικά διχωρίσιμα γιαυτό το λόγο προτιμούμε την μέθοδο ταξινόμησης με συναρτήσεις δυναμικού.

4 Μονάδες 2

Κατασκευάστε ένα πολυεπίπεδο δίκτυο το οποίο να αποτελείται από δύο επίπεδα νευρώνων. Το πρώτο επίπεδο έχει τρείς νευρώνες, ενώ το δεύτερο επίπεδο (το επίπεδο εξόδου) έχει μόνον δύο νευρώνες. Οι νευρώνες διαθέτουν τον ακόλουθο μη γραμμικό τελεστή:

$$f(x) = \frac{x}{1+|x|}$$

Τοποθετείστε στην είσοδο του δικτύου τα ψηφία του αριθμού μητρώου σας και επαναπροσδιορίστε τους συντελεστές βαρύτητας των νευρώνων με την μέθοδο της Hebbian εκπαίδευσης.

Λύση:

Ονομάζω τους τρείς νευρώνες του πρώτου επιπέδου N_1, N_2, N_3 και του επιπέδου εξόδου N_4, N_5.

Επιλέγω τυχαία (μικρούς αριθμούς κοντά στην τιμή μηδέν) τους συντελεστές βαρύτητας των νευρώνων:

$$\mathbf{w}_1 = (-0.1, 0.1, 0.1, 0.1)$$

$$\mathbf{w}_2 = (0.1, -0.1, 0.1, 0.1)$$

$$\mathbf{w}_3 = (0.1, 0.1, -0.1, 0.1)$$

$$\mathbf{w}_4 = (-0.1, 0.1, 0.1)$$

$$\mathbf{w}_5 = (0.1, -0.1, 0.1)$$

Τοποθετώ στην είσοδο τα ψηφία του αριθμού μητρώου μου, έστω ότι είναι ο 1234.

Οι έξοδοι του πρώτου επιπέδου νευρώνων θα είναι

$$f_1((1,2,3,4)\mathbf{w}_1^T) = f_1(0.8) = 0.4444$$

$$f_2((1,2,3,4)\mathbf{w}_2^T) = f_1(0.6) = 0.375$$

$$f_3((1,2,3,4)\mathbf{w}_3^T) = f_1(0.4) = 0.2857$$

Η Hebbian εκπαίδευση επαναπροσδιορίζει τους συντελεστές βαρύτητας με την βοήθεια της σχέσης $w_{ij} = w_{ij} + a o_i o_j$ όπου o_i και o_j είναι οι έξοδοι των νευρώνων που συνδέει η σύναψη που έχει συντελεστή βαρύτητας τον w_{ij}. Τοποθετώ a=1. Συνεπώς οι νέοι συντελεστές βαρύτητας γιά το πρώπο επίπεδο νευρώνων είναι οι ακόλουθοι:

$$\mathbf{w}_1 = (-0.1 + (1) * (0.4444), 0.1 + (2) * (0.4444), 0.1 + (3) * (0.4444), 0.1 + (4) * (0.4444)$$

$$\mathbf{w}_2 = (0.1 + (1) * (0.375), -0.1 + (2) * (0.375), 0.1 + (3) * (0.275), 0.1 + (4) * (0.375)$$

$$\mathbf{w}_3 = (0.1 + (1) * (0.2857), 0.1 + (2) * (0.2857), -0.1 + (3) * (0.2857), 0.1 + (4) * (0.2857)$$

Η ίδια διαδικασία επαναλαμβάνεται και για τους δύο νευρώνες του επιπέδου εξόδου.

Επιλέγω τυχαία (μικρούς αριθμούς κοντά στην τιμή μηδέν) τους συντελεστές βαρύτητας των νευρώνων:

$$\mathbf{w}_4 = (-0.1, 0.1, 0.1)$$

$$\mathbf{w}_5 = (0.1, -0.1, 0.1)$$

Τοποθετώ στην είσοδο τις τιμές εξόδου των νευρώνων του πρώτου επιπέδου

Οι έξοδοι του δεύτερου επιπέδου νευρώνων θα είναι

$$f_4((0.4444, 0.375, 0.2857)\mathbf{w}_4^T) = f_1(0.02163) = 0.02117$$

$$f_5((0.4444, 0.375, 0.2857)\mathbf{w}_5^T) = f_1(0.03569) = 0.07138$$

Οι νέοι συντελεστές βαρύτητας γιά το πρώπο επίπεδο νευρώνων είναι οι ακόλουθοι:

$$\mathbf{w}_4 = (-0.1 + (0.4444) * (0.02117), 0.1 + (0.375) * (0.02117), 0.1 + (0.2857) * (0.02117))$$

$$\mathbf{w}_5 = (-0.1 + (0.4444) * (0.07138), 0.1 + (0.375) * (0.07138), 0.1 + (0.2857) * (0.07138))$$