ΑΝΑΓΝΩΡΙΣΗ ΠΡΟΤΥΠΩΝ Ι
Τμήμα Ηλεκτρολόγων Μηχανικών και Τεχνολογίας Υπολογιστών
Εξεταστική περίοδος Ιούλιος 2007

Άσκηση 1. Μονάδες 5

Για την διάγνωση της ασθένειας Χ υπάρχουν δύο εξετάσεις η πρώτη κοστίζει 5 Ευρώ και η δεύτερη 100 Ευρώ. Η πρώτη εξέταση παράγει έναν αριθμό ο οποίος στην περίπτωση που ο εξεταζόμενος είναι υγιής έχει συνάρτηση πυκνότητας πιθανότητας που δίνεται από την δευτεροβάθμια εξίσωση για την οποία ισχύει,

$$f(x \mid \text{ΥΓΙΗΣ}) = \begin{cases} ax^2 + bx + c, & 0 \le x \le 8 \\ 0, & \alpha\lambda\lambda o\acute{u} \end{cases} \quad \text{με } \alpha<0.$$

Στην περίπτωση κατά την οποία ο εξεταζόμενος είναι φορέας της ασθένειας έχει υπολογιστεί ότι η συνάρτηση πυκνότητας πιθανότητας της μέτρησης x ακολουθεί την τριγωνική κατανομή:

$$f(x \mid \text{ΦΟΡΕΑΣ}) = \begin{cases} dx + g, & 3 \le x \le 10 \\ 0, & \alpha\lambda\lambda o\acute{u} \end{cases} \quad \text{με } d>0.$$

Η δεύτερη διαγνωστική μέθοδος αναγνωρίζει με μηδενικό σφάλμα την ύπαρξη της ασθένειας Χ στον εξεταζόμενο.

Λαμβάνοντας υπόψιν ότι ο πληθυσμός που θα εξετάσετε είναι 100000 άνθρωποι από τους οποίους εκτιμάται ότι υπάρχουν 20000 φορείς, κάντε τις υποθέσεις που εσείς θεωρείτε λογικές και απαραίτητες και εκτελέστε τους ακόλουθους υπολογισμούς:

Α. Προτείνετε διαγνωστική διαδικασία η οποία να ελαχιστοποιεί το διαγνωστικό σφάλμα χρησιμοποιώντας μονάχα την πρώτη εξέταση. Υπολογίστε το συνολικό κόστος εξέτασης όλου του πληθυσμού. Πόσες λανθασμένες διαγνώσεις θα έχετε;

Β. Προτείνετε διαγνωστική διαδικασία η οποία να ελαχιστοποιεί το κόστος εξέτασης. Υπολογίστε το συνολικό κόστος. Πόσες λανθασμένες διαγνώσεις θα έχετε;

Γ. Έστω ότι διαθέτετε 2 εκατ. Ευρώ για να εκτελέσετε διαγνωστικές εξετάσεις. Τι θα κάνετε για να ελαχιστοποιήσετε τον αριθμό των λανθασμένων διαγνώσεων; Πόσες λανθασμένες διαγνώσεις θα έχετε;

Λύση Ασκησης 1

Για να λύσω την άσκηση χρειάζεται να υπολογίσω τις σταθερές στις δεσμευμένες σ.π.π. Μου δίνεται μόνο η μορφή τους και το διάστημα που οι συναρτήσεις έχουν μη μηδενική τιμή.

Ο βέλτιστος υπολογισμός της συνάρτησης πυκνότητας πιθανότητας (σ.π.π.) πρέπει να γίνει με αμερόληπτο τρόπο, δηλαδή οι σταθερές *a,b,c,d,g* πρέπει να υπολογιστούν έτσι ώστε να μεγιστοποιείται η εντροπία της σ.π.π. και να ικανοποιούνται οι αξιωματικοί περιορισμοί (μη αρνητικές τιμές και το ολοκλήρωμα να είναι ίσο με 1).

Το πρόβλημα έτσι όπως τίθεται δεν έχει εύκολη μαθηματική λύση, αλλά γνωρίζω ότι η ομοιόμορφη σ.π.π μεγιστοποιεί την εντροπία, όταν ο μόνος περιορισμός που γνωρίζω είναι τα διαστήματα τιμών της σ.π.π., όπως συμβαίνει και στην άσκηση. Συνεπώς θα ορίσω τις παραμέτρους έτσι ώστε οι δοσμένες συναρτήσεις να προσεγγίζουν την ομοιόμορφη σ.π.π.

Δηλαδή

$$f(x \mid \text{ΥΓΙΗΣ}) = \begin{cases} -10^{-1000}x^2 + 1/8, & 0 \le x \le 8 \\ 0, & \alpha\lambda\lambda o\acute{u} \end{cases} \quad (\alpha<0).$$

και

$$f(x \mid \text{ΦΟΡΕΑΣ}) = \begin{cases} 10^{-1000}x + 1/7, & 3 \le x \le 10 \\ 0, & \alpha\lambda\lambda o\acute{u} \end{cases} \quad (d>0).$$

Ερώτημα Α.

Δεν μου δίνεται περιορισμός χρημάτων. Συνεπώς κάνω όσες εξετάσεις θέλω από το πρώτο είδος (Ε1).

Αρχικά έχω να εξετάσω πληθυσμό 100000, 80000 υγιείς και 20000 ασθενείς. Κάνω σε όλους την Ε1. Οσοι δώσουν τιμή στο διάστημα [0,3) είναι σίγουρα υγιείς ενώ όσοι δώσουν τιμές στο διάστημα (8,10] είναι σίγουρα φορείς. Μετά την πρώτη Ε1 αναμένεται να βρεθούν

$$80000 * \int_0^3 -10^{-1000} x^2 + 1/8 dx \simeq 80000 * (3/8) = 30000 \text{ υγιείς, και}$$

$$20000 * \int_8^{10} 10^{-1000} x + 1/7 dx \simeq 20000 * (2/7) = 5714 \text{ ασθενείς.}$$

Μένει ένας πληθυσμός 100000-30000-5714 = 64286, (80000-30000=50000 υγιείς, 20000-5714=14286 φορείς) για τους οποίους δεν μπορώ να αποφασίσω με σιγουριά γιατί δίνουν τιμές εξέτασης στο διάστημα [3,8].
Αν υποθέσω ότι οι εξετάσεις στον ίδιο άνθρωπο είναι μεταξύ τους στοχαστικά ανεξάρτητες τότε θα επαναλάβω την Ε1 στους 64286 ανθρώπους. Οπότε θα έχω με το ίδιο σκεπτικό τα ακόλουθα αποτελέσματα

$$50000 * \int_0^3 -10^{-1000} x^2 + 1/8 dx \simeq 50000 * (3/8) = 18750 \text{ υγιείς, και}$$

$$14286 * \int_8^{10} 10^{-1000} x + 1/7 dx \simeq 14286 * (2/7) = 4081 \text{ ασθενείς.}$$

Μένει ένας πληθυσμός 64286-18750-4081=41455, (50000-18750=31250 υγιείς, 14286-4081=10250 φορείς) για τους οποίους δεν μπορώ να αποφασίσω με σιγουριά.
Καταννοώ λοιπόν ότι αν επαναλαμβάνω συνέχεια την εξέταση Ε1 στον πληθυσμό για τον οποίο δεν μπορώ να αποφασίσω με σιγουριά, τότε με πεπερασμένο αριθμό εξετάσεων Ε1 θα διαγνώσω για τους 100000 ανθρώπους με σφάλμα 0% αν είναι υγιείς ή φορείς.
Το συνολικό κόστος εξέτασης προκύπτει από αναδρομικό αλγόριθμο ώς εξής:
Βημα0. Αρχικός πληθυσμός που αποτελείται από Υ(0) υγιείς και Φ(0) φορείς. t=0, K(0)=0.
Βημα1. t=t+1. Συνολικό Κόστος εξέτασης K(t)= K(t-1) + (Υ(t-1)+Φ(t-1))*5 Ευρώ.
Υ(t)=Υ(t-1)- Υ(t-1)(3/8)= Υ(t-1)5/8
Φ(t)=Φ(t-1)- Φ(t-1)(2/7)= Φ(t-1)5/7
Βημα2. Αν Φ(t) ή Υ(t) είναι 0 τότε ο αλγόριθμος τερματίζει και το K(t) περιέχει το συνολικό κόστος της εξέτασης. Διαφορετικά επαναλαμβάνεται το βήμα 1.

Εναλλακτική λύση
Αν υποθέσω ότι η επανάληψη της Ε1 στον άνθρωπο δίνει την ίδια αριθμητική τιμή, τότε η λύση είναι πολύ απλή. Το συνολικό κόστος της μοναδικής εξέτασης που κάνω σε όλον τον πληθυσμό είναι 100000*5 = 500000 Ευρώ. Στο διάστημα τιμών [3,8] βρίσκονται 31250 υγιείς και 10250 ασθενείς και οι σ.π.π. είναι πρακτικά ομοιόμορφες. Για να έχω ελάχιστο σφάλμα θα αποφασίζω στο διάστημα [3,8] ότι ο εξεταζόμενος είναι υγιής. Το συνολικό σφάλμα του συστήματος της μοναδικής εξέτασης Ε1 θα δίνει πιθανότητα σφάλματος 14286/100000% δηλ. 14.286%.

Ερώτημα Β.
Δεν κάνω εξέταση σε κανένα άνθρωπο και αποφασίζω ότι όλοι είναι υγιείς. Συνεπώς το κόστος της εξέτασης είναι 0 Ευρώ και το αντίστοιχο σφάλμα του συστήματος απόφασης είναι 20000/100000=20%.

Ερώτημα Γ.
Ανάλογα με την υπόθεση που έχω κάνει στο ερώτημα Α θα έχω και την αντίστοιχη λύση στο ερώτημα Γ.
Υπόθεση: Οι εξετάσεις στον ίδιο άνθρωπο είναι μεταξύ τους στοχαστικά ανεξάρτητες. Εφαρμόζω τον αλγόριθμο του ερωτήματος Α και βλέπω ότι θα χρειαστώ 1416200 ευρώ για να πετύχω σφάλμα 0% για τους 100000 ανθρώπους. Δηλαδή θα μου περισσέψουν και χρήματα.

Υπόθεση: Η επανάληψη της Ε1 στον άνθρωπο δίνει την ίδια αριθμητική τιμή.
Κάνω την Ε1 στους 100000. Κόστος 500000 Ευρώ και τα 1500000 Ευρώ τα χρησιμοποιώ για την εξέταση Ε2. Εχω δυνατότητα να εξετάσω 1500000/100=15000 ανθρώπους.
Από την Ε1 θα έχω στο διάστημα τιμών [3,8] 50000 υγιείς (50000/64286=77.77%), και 14286 φορείς (14286/64286=22.23%). Επιλέγω τυχαία 15000 ανθρώπους και κάνω την εξέταση Ε2. Θα διαγνώσω 15000*0.7777=11665 υγιείς, και 14286*0.2223=3175 φορείς. Αποφασίζω πάλι ότι στο διάστημα τιμών [3,8] όλοι είναι υγιείς. Το σφάλμα που θα έχω είναι ότι οι φορείς που δεν έχουν εξεταστεί με την Ε2 και είναι στο διάστημα [3,8] είναι 14286-3175=11111. Το σφάλμα της διαδικασίας εξέτασης είναι:
11111/100000 = 11.11%.

Ασκηση 2. Μονάδες 3
Θέλετε να κατασκευάσετε σύστημα αναγνώρισης προτύπων δύο κατηγοριών που χρησιμοποιεί το κριτήριο ταξινόμησης του πλησιέστερου γείτονα και συνάρτηση απόστασης την γνωστή σαν city-block:

$$d(x, y) = \sum_{i=1}^{N} | x_i - y_i |$$

Υπολογίστε τα βέλτιστα εικονικά πρωτότυπα (τα προτυπα που ελαχιστοποιούν την αθροιστική απόσταση) όταν σας δίνονται τα ακόλουθα παραδείγματα:

Κατηγορία 1	(-2,1)	(-0.5,1.5)	(2.5,0.5)	(3,-1.5)		
Κατηγορία 2	(-1,-1)	(0.5,0)	(1,1)	(1.5,0.5)	(2,-1)	(2,-2)

Λύση Ασκησης 2

Το βέλτιστο εικονικό πρωτότυπο υπολογίζεται από την ελαχιστοποίηση του συνάρτησης

$$\sum_{j=1}^{M} d(y, x_j) = \sum_{j=1}^{M}\sum_{i=1}^{N} | y_i - x_{ji} | = \sum_{i=1}^{N}(\sum_{j=1}^{M} | y_i - x_{ji} |)$$

Όπου y είναι το βέλτιστο εικονικό πρωτότυπο, x_j είναι τα πρότυπα, M είναι ο αριθμός των προτύπων και N είναι η διάσταση των προτύπων.
Από την δεύτερη σχέση (άθροισμα M θετικών όρων) βλέπω ότι μπορώ να λύσω το πρόβλημα ελαχιστοποίησης για κάθε y_i ανεξάρτητα από τα υπόλοιπα. Συνεπώς έχω να λύσω το πρόβλημα εύρεσης του ελάχιστου της $O(y_i)$:

$$O(y_i) = \sum_{j=1}^{M} | y_i - x_{ji} |$$

Η συνάρτηση |.| είναι συνεχής οπότε και η O(.) είναι συνεχής σαν άθροισμα συνεχών συναρτήσεων. Παρατηρώ επίσης ότι για να απαλείψω την απόλυτη τιμή θα πρέπει να γνωρίζω την σχετική θέση του y_i ως προς τις μετρήσεις x_{ji}. Αν τοποθετήσω τις M μετρήσεις κατά αύξουσα αριθμητική σειρά θα έχω:

y_i	$-\infty$	X_{1i}	X_{2i}	...	X_{M-1i}	X_{Mi} $+\infty$
$O(y_i)$	$-My_i + \sum_{j=1}^{M} x_{ji}$	$-(M-2)y_i + ...$			$(M-2)y_i - ...$	$My_i - \sum_{j=1}^{M} x_{ji}$

Βλέπω ότι αρχικά η συνάρτηση $O(y_i)$ είναι φθίνουσα (ο συντελεστής του y_i είναι αρνητικός), η κλίση της ευθείας στα διαστήματα ελαττώνεται ($-M$, $-(M-2)$, ...) και στην συνέχεια η κλίση της συνάρτησης $O(.)$ γίνεται θετική με την κλίση συνεχώς να αυξάνει. Συνεπώς το ελάχιστο της $O(.)$ βρίσκεται στο μέσο της διατεταγμένης τοποθέτησης των προτύπων για κάθε διάσταση και κατηγορία ανεξάρτητα. Δεν προκύπτουν μοναδικές λύσεις διότι σε μερικά διαστήματα τιμών το ελάχιστο της $O(.)$ είναι ένα οριζόντιο τμήμα ευθείας. Η λύση λοιπός έχει ως εξής:
Τοποθετώ σε κάθε διάσταση και κατηγορία τα παραδείγματα κατά αύξουσα αριθμητική σειρά και πηγαίνω στο μεσαίο διάστημα και διαλέγω μια τιμή για το εικονικό πρότυπο.
Συνεπώς η άσκηση αυτή μπορεί να λυθεί χωρίς να κάνετε καμμία αριθμητική πράξη.

Άσκηση 3. Μονάδες 4

Να δώσετε τις σχέσεις επαναπροσδιορισμού των συντελεστών βαρύτητας των συνάψεων νευρωνικού δικτύου ακτινικών συναρτήσεων και να περιγράψετε το τρόπο με τον οποίο θα το εκπαιδεύσετε με παραδείγματα χρησιμοποιώντας τον αλγόριθμο οπισθοδρομικής διάδοσης του σφάλματος.

Ο μη-γραμμικός πυρήνας στο κρυφό επίπεδο δίνεται από την σχέση

$$d(w, y) = \sum_{i=1}^{N} | w_i - y_i |$$

Λύση Ασκησης 3

Από τις σημειώσεις του μαθήματος (σελ.123) γνωρίζω τις σχέσεις επαναπροσδιορισμού των συντελεστών βαρύτητας των συνάψεων νευρωνικού δικτύου. Το νευρωνικό δίκτυο που δίνεται από την άσκηση περιέχει νευρώνες που δεν ταιριάζουν ακριβώς με το μοντέλο που χρειάζεται ο αλγόριθμος οπισθοδρομικής διάδοσης του σφάλματος. Στο πρώτο επίπεδο υπάρχει μόνο ένας κατά τμήματα γραμμικός τελεστής, η συνάρτηση απόλυτη τιμή, και στην συνέχεια υπάρχει ο γραμμικός συνδυασμός τους. Σε κάθε έξοδο του νευρωνικού δικτύου η αριθμητική του τιμή σαν συνάρτηση της εισόδου και και των συντελεστών βαρύτητας των συνάψεων δίνεται από την σχέση:

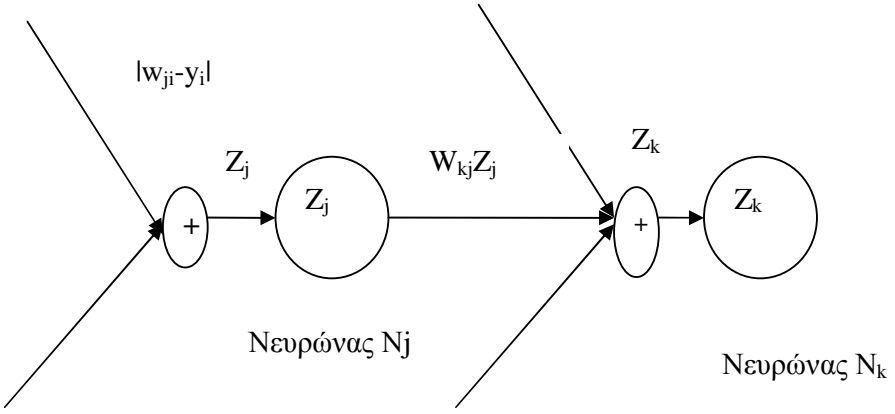
$$O(y) = \sum_{j=1}^{M} W_j \sum_{i=1}^{N} | w_{ji} - y_i | \qquad (1)$$

Για να μπορέσουμε να εφαρμόσουμε τις σχέσεις επαναπροσδιορισμού των συντελεστών βαρύτητας συνάψεων νευρωνικού δικτύου, πρέπει να περιγράψουμε της σχέση (1) σαν μια σύνδεση νευρώνων. Ο κάθε νευρώνας θα πρέπει να αποτελείται απο δύο σειριακά συνδεδεμένους τελεστές. Ο πρώτος εκτελεί έναν μετασχηματισμό ο οποίος συνδέει τους συνετελεστές βαρύτητας των συνάψεων και ο δεύτερος εκτελεί έναν μη γραμμικό μετασχηματισμό της εσωτερικής κατάστασης του νευρώνα (βλ. Σχήμα 4.7, σελ 123).

Η άσκηση μπορεί να λυθεί με πολλούς τρόπους ανάλογα με τις θεωρήσεις που θα κάνουμε σε αυτό το σημείο. Στην συνέχεια περιγράφεται μια από τις πλέον απλές λύσεις (όχι η απλούστερη).

Επειδή έχω δύο ενφωλιασμένα αθροίσματα καταλαβαίνω ότι πρέπει να κατασκευάσω δύο επίπεδα νευρώνων τα οποία θα μου δίνουν την ίδια ακριβώς συνάρτηση μεταφοράς με την (1) και θα αποτελούνται από νευρώνες που θα έχουν τα χαρακτηριστικά που περιγράφηκαν.

Εχω λοιπόν το ακόλουθο νευρωνικό δίκτυο



Νευρώνας Nj: Μετασχηματισμός των δεδομένων εισόδου με τους νευρώνες $|w_{ji}-y_i|$

Μη-γραμμικός μετασχηματισμός: $f(Z_j) = Z_j$ (δεν υπάρχει μετασχ)

Νευρώνας N_k: Μετασχηματισμός των δεδομένων εισόδου με τους νευρώνες $W_{kj}Z_j$
Μη-γραμμικός μετασχηματισμός: $f(Z_k) = Z_k$ (δεν υπάρχει)

Εφόσον η τυποποίηση που έκανα ταιριάζει με αυτήν του σχήματος (βλ. Σχήμα 4.7, σελ 123), δεν μένει παρά να εφαρμόσω τις σχέσεις επαναπροσδιορισμού, ως εξής:

Ορίζω την ίδια συνάρτηση σφάλματος (4.34), οπότε η σχέση επαναπροσδιορισμού των w (4.38) γίνεται ως εξής:

Νευρώνες εξόδου (4.38): $\Delta W_{kj} = n\delta_k \frac{\partial Z_k}{\partial w_{kj}} = n\delta_k Z_j$

Κρυφοί νευρώνες (4.38): $\Delta w_{ji} = n\delta_j \frac{\partial Z_j}{\partial w_{ji}} = n\delta_j \begin{cases} 1, & w_{ji} \geq y_i \\ -1, & w_{ji} < y_i \end{cases}$

Ο υπολογισμός των δ γίνεται ως εξής:

Νευρώνες εξόδου (4.39): $\delta_k = (\beta_k - Z_k)\frac{\partial f(Z_k)}{\partial Z_k} = (\beta_k - Z_k)\frac{\partial Z_k}{\partial Z_k} = (\beta_k - Z_k)$

Κρυφοί νευρώνες (4.41): $\delta_j = \frac{\partial Z_j}{\partial Z_j}\sum_{k=1}^{K}(-\frac{\partial \Sigma\phi\acute{\alpha}\lambda\mu\alpha}{\partial Z_k})\frac{\partial Z_k}{\partial Z_j} = \sum_{k=1}^{K}\delta_k W_{kj}$