

Ασκήσεις μελέτης της 7^{ης} διάλεξης

7.1. (α) Βάσει των δεδομένων εκπαίδευσης του διπλανού πίνακα, η εντροπία της κατηγορίας C είναι:

$$\underline{X_H(C)} = 1 \quad \underline{H(C)} = 0 \quad \underline{H(C)} = \frac{1}{2}$$

X	Y	Z	C
0	1	0	θετικό
0	1	1	θετικό
0	1	0	θετικό
1	0	1	θετικό
1	1	0	αρνητικό
0	0	1	αρνητικό
0	0	0	αρνητικό
0	0	1	αρνητικό

Απάντηση: $P(C=\text{θετικό}) = P(C=\text{αρνητικό}) = \frac{1}{2}$. Έχουμε δύο ισοπίθανα ενδεχόμενα $C = \text{θετικό}$ και $C = \text{αρνητικό}$ και άρα μέγιστη εντροπία (αβεβαιότητα), που για δύο ενδεχόμενα είναι ίση με 1. Το ίδιο αποτέλεσμα προκύπτει από τον ορισμό της εντροπίας, κάνοντας τις πράξεις.

β) Βάσει των δεδομένων του πίνακα του σκέλους (α), υψηλότερο είναι το κέρδος πληροφορίας (δεν χρειάζονται πράξεις):

$$\underline{IG(C, X)} \quad \underline{X_IG(C, Y)} \quad \underline{IG(C, Z)}$$

Απάντηση: Τα δεδομένα εκπαίδευσης δείχνουν ότι αν μάθουμε πως $Y = 1$, είναι πολύ πιθανό (πιθανότητα $\frac{3}{4}$) ότι $C = \text{θετικό}$ και αν μάθουμε πως $Y = 0$, είναι πολύ πιθανό (πιθανότητα $\frac{3}{4}$) ότι $C = \text{αρνητικό}$. Επομένως, η γνώση της τιμής της ιδιότητας Y μειώνει την εντροπία (αβεβαιότητα για την τιμή) της C , εντροπία που αρχικά ήταν μέγιστη. Το ίδιο συμπέρασμα προκύπτει υπολογίζοντας το $IG(C, Y)$.

Αντίθετα, τα δεδομένα εκπαίδευσης δείχνουν ότι αν μάθουμε πως $X = 1$, η πιθανότητα να έχουμε $C = \text{θετικό}$ παραμένει $\frac{1}{2}$ και ίση με την πιθανότητα να έχουμε $C = \text{αρνητικό}$ και αν μάθουμε ότι $X = 0$, η πιθανότητα να έχουμε $C = \text{θετικό}$ παραμένει $\frac{1}{2}$ και ίση με την πιθανότητα να έχουμε $C = \text{αρνητικό}$. Επομένως, όποια κι αν είναι η τιμή της X , εξακολουθούμε να έχουμε δύο ισοπίθανα ενδεχόμενα $C = \text{θετικό}$ και $C = \text{αρνητικό}$ και, επομένως, μέγιστη εντροπία (αβεβαιότητα). Άρα η γνώση της τιμής της X δεν μειώνει καθόλου την εντροπία (αβεβαιότητα για την τιμή της) C , επομένως $IG(C, X) = 0$. Το ίδιο συμπέρασμα προκύπτει υπολογίζοντας το $IG(C, X)$.

Ομοίως, $IG(C, Z) = 0$.

7.2. Υπολογίστε την εντροπία στις περιπτώσεις που αναφέρει η διαφάνεια 24 (μηνύματα ηλεκτρονικού ταχυδρομείου). Υπόδειξη: Για $P(C) \rightarrow 0$, χρησιμοποιήστε τον κανόνα De L'Hôpital.

Απάντηση:

α) Όταν τα παραδείγματα εκπαίδευσης περιλαμβάνουν 200 ανεπιθύμητα και 600 επιθυμητά μηνύματα, έχουμε:

$$P(C = 1) = \frac{200}{800} = \frac{1}{4}, \quad P(C = 0) = \frac{600}{800} = \frac{3}{4},$$

$$\log_2 \frac{1}{4} = -2, \quad \log_2 \frac{3}{4} = \log_2 3 - \log_2 4 = 1.585 - 2 = -0.415$$

$$H(C) = \frac{1}{4} \cdot 2 + \frac{3}{4} \cdot 0.415 = 0.811$$

β) Όταν τα παραδείγματα εκπαίδευσης περιλαμβάνουν 400 ανεπιθύμητα και 400 επιθυμητά μηνύματα, έχουμε:

$$P(C = 1) = \frac{400}{800} = \frac{1}{2}, P(C = 0) = \frac{400}{800} = \frac{1}{2}, \log_2 \frac{1}{2} = -1,$$

$$H(C) = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 1 = 1$$

γ) Όταν τα παραδείγματα εκπαίδευσης τείνουν να είναι όλα ανεπιθύμητα, έχουμε:

$$P(C = 1) = 1, P(C = 0) = 0, \log_2 P(C = 1) = 0, \log_2 P(C = 0) \rightarrow -\infty$$

$$\begin{aligned} H(C) &= -P(C = 1) \cdot \log_2 P(C = 1) - P(C = 0) \cdot \log_2 P(C = 0) = \\ &= -1 \cdot 0 - \frac{\log_2 P(C = 0)}{\frac{1}{P(C = 0)}} \end{aligned}$$

Χρησιμοποιώντας τον κανόνα De L'Hôpital για τον δεύτερο προσθετέο, θεωρώντας ότι $P(C = 0) \rightarrow 0^+$, καταλήγουμε:

$$H(C) \rightarrow -\frac{\frac{1}{P(C = 0) \ln 2}}{-\frac{1}{P(C = 0)^2}} = \frac{P(C = 0)}{\ln 2} = 0$$

Ομοίως, όταν όλα τα παραδείγματα εκπαίδευσης είναι επιθυμητά, έχουμε πάλι $H(C) = 0$.

7.3. (α) Εξηγήστε γιατί οι αλγόριθμοι επιβλεπόμενης μάθησης πρέπει να αξιολογούνται σε διαφορετικά δεδομένα από αυτά με τα οποία εκπαιδεύτηκαν.

Απάντηση: Αν αξιολογούνταν στα δεδομένα εκπαίδευσης, ένας αλγόριθμος που θα απομνημόνευε τις σωστές απαντήσεις των παραδειγμάτων εκπαίδευσης και θα απαντούσε τυχαία σε κάθε άλλη περίπτωση θα πετύχαινε ακρίβεια 100% στην αξιολόγηση, χωρίς αυτή η επίδοση να είναι ενδεικτική του πόσο καλά θα τα πήγαινε σε διαφορετικά δεδομένα αξιολόγησης, αφού τότε θα απαντούσε τυχαία. Γενικότερα, υπάρχει ο κίνδυνος το μοντέλο απόφασης που παράγεται να είναι υπερ-εξειδικευμένο στα δεδομένα εκπαίδευσης αξιολόγησης, με αποτέλεσμα να επιτυγχάνεται υψηλή ακρίβεια στα δεδομένα εκπαίδευσης, που δεν είναι ενδεικτική της ακρίβειας που επιτυγχάνεται σε διαφορετικά δεδομένα.

(β) Ένας ερευνητής υπέβαλε σε ένα επιστημονικό συνέδριο εργασία του στην οποία περιέγραφε ένα σύστημα κατάταξης κειμένων που χρησιμοποιούσε επιβλεπόμενη μηχανική μάθηση. Στην εργασία περιέγραφε, μεταξύ άλλων, πειράματα στα οποία δοκίμασε πολλά διαφορετικά σύνολα ιδιοτήτων. Για κάθε σύνολο ιδιοτήτων, είχε εκπαιδεύσει το σύστημα σε ένα σύνολο κειμένων εκπαίδευσης (το ίδιο για όλα τα σύνολα ιδιοτήτων) και τον είχε αξιολογήσει σε ένα εντελώς διαφορετικό σύνολο κειμένων αξιολόγησης (το ίδιο για όλα τα σύνολα ιδιοτήτων). Στην εργασία παρέθετε τα αποτελέσματα της αξιολόγησης για κάθε διαφορετικό σύνολο ιδιοτήτων, από τα οποία προέκυπτε το καλύτερο σύνολο ιδιοτήτων και η αντίστοιχη καλύτερη επίδοση του συστήματος στο σύνολο αξιολόγησης. Ωστόσο, οι κριτές του συνεδρίου απέρριψαν την εργασία λέγοντας ότι μέρος της εκπαίδευσης είχε γίνει στα δεδομένα αξιολόγησης. Είχαν δίκιο οι κριτές; Εξηγήστε γιατί. Αν πιστεύετε ότι είχαν δίκιο, εξηγήστε επίσης τι θα έπρεπε να κάνει ο ερευνητής για να αντιμετωπίσει το πρόβλημα.

Απάντηση: Είχαν δίκιο, γιατί ο ερευνητής επέλεξε τις ιδιότητες που οδηγούσαν στα καλύτερα αποτελέσματα αξιολόγησης. Ουσιαστικά, δηλαδή, χρησιμοποίησε τα δεδομένα αξιολόγησης για

την επιλογή ιδιοτήτων, η οποία αποτελεί μέρος της εκπαίδευσης. Υπάρχει ο κίνδυνος να επέλεξε έτσι ιδιότητες που οδηγούν σε καλά αποτελέσματα στο συγκεκριμένο σύνολο αξιολόγησης και μόνο (πρόβλημα υπερ-εφαρμογής). Θα έπρεπε να είχε χρησιμοποιήσει ένα ξεχωριστό σύνολο δεδομένων επικύρωσης. Για κάθε σύνολο ιδιοτήτων, θα έπρεπε να είχε εκπαιδεύσει το σύστημα με τα δεδομένα εκπαίδευσης και να είχε μετρήσει την επίδοσή του στα δεδομένα επικύρωσης, ώστε να επιλέξει το σύνολο ιδιοτήτων με την καλύτερη επίδοση στα δεδομένα επικύρωσης. Κατόπιν, θα έπρεπε να είχε αξιολογήσει το σύστημα με το επιλεγμένο σύνολο ιδιοτήτων στα δεδομένα αξιολόγησης.