Ε.Μ.Π., Σχολή Η.Μ.& Μ.Υ.	**Αναγνώριση Προτύπων**	Ακαδ.Ετος 2025-26
Ημερ/νία: 11-11-2025	**Σύνολο Αναλυτικών Ασκήσεων**	Παραδοτέο: 25-11-2025

Οι αναλυτικές σειρές ασκήσεων είναι ατομικές, και οι λύσεις που θα δώσετε πρέπει να αντιπροσωπεύουν μόνο την προσωπική σας εργασία. Εξηγήστε επαρκώς την εργασία σας. Αν χρησιμοποιήσετε κάποια άλλη πηγή εκτός των σημειώσεων για τη λύση σας, πρέπει να το αναφέρετε. Η παράδοση των λύσεων των αναλυτικών ασκήσεων της σειράς αυτής θα γίνει ηλεκτρονικά στο helios και θα πρέπει να την υποβάλετε ως ένα ενιαίο αρχείο PDF με το εξής filename format, χρησιμοποιώντας μόνο λατινικούς χαρακτήρες: pr23_hwk_AM_FirstnameLastname.pdf, όπου AM είναι ο 8-ψήφιος αριθμός μητρώου σας. Σκαναρισμένες χειρόγραφες λύσεις επιτρέπονται, αρκεί να είναι καθαρογραμμένες και ευανάγνωστες. Επίσης, στην πρώτη σελίδα των λύσεων θα αναγράφετε το ονοματεπώνυμο, AM, και email address σας. Απαγορεύεται αυστηρά η ανάρτηση των λύσεων σε οποιαδήποτε πλατφόρμα καθώς το περιεχόμενο της άσκησης αποτελεί πνευματική ιδιοκτησία των διδασκόντων και της ομάδας υποστήριξης του μαθήματος. Οι αναλυτικές σειρές ασκήσεων είναι ατομικές, και οι λύσεις που θα δώσετε πρέπει να αντιπροσωπεύουν μόνο την προσωπική σας εργασία.

Υλικό για Ανάγνωση: Βιβλία: [1], [2], [3], [4], [5]
Διαφάνειες διαλέξεων μαθήματος

Αναλυτικές Ασκήσεις

---

1 Ταξινόμηση στη μία διάσταση

Σετ Δεδομένων

Δίδεται ένα συνθετικό σύνολο δεδομένων με μονοδιάστατο χαρακτηριστικό x που αντιπροσωπεύει κάποια μέτρηση, με δύο κλάσεις (A και B). Η κλάση A έχει σημεία $x = \{1, 3, 5\}$ και η κλάση B έχει σημεία $x = \{2, 4, 6\}$.

1. Ταξινομητής Κοντινότερου Γείτονα

(a) Υπολογίστε το όριο απόφασης για έναν ταξινομητή Κοντινότερου Γείτονα 1-NN και 3-NN και σχεδιάστε τα σημεία των κλάσεων στην πραγματική ευθεία με βάση τους δύο αυτούς ταξινομητές.

2. Αλγόριθμος Perceptron

(a) Αρχικοποιήστε το perceptron με βάρη $w = 0$ και πόλωση $b = 0$.

(b) Εκτελέστε τον αλγόριθμο εκμάθησης perceptron στο σετ δεδομένων και βήμα εκμάθησης 1.

(c) Προσδιορίστε τα βάρη και την πόλωση μετά από μια εποχή. Πως ταξινομούνται πλέον τα δεδομένα; Μπορεί το perceptron να τα διαχωρίσει;

(d) Εκτελέστε μια ακόμα εποχή για να δείξετε πειραματικά το παραπάνω.

3. Εκτίμηση Μέγιστης Πιθανοφάνειας για Γκαουσιανές Κατανομές

(a) Υποθέστε ότι οι κατανομές και για τις δύο κλάσεις είναι Γκαουσιανές. Υπολογίστε τον μέσο όρο και την διακύμανση για κάθε κλάση χρησιμοποιώντας την Εκτίμηση Μέγιστης Πιθανοφάνειας. Δείξτε αναλυτικά τη διαδικασία εξαγωγής των τύπων.

4. Σφάλμα Ταξινόμησης Bayes

Χρησιμοποιώντας τις Γκαουσιανές κατανομές από την Ερώτημα 3, υπολογίστε την πιθανότητα λανθασμένης ταξινόμησης για ένα νέο σημείο στην πραγματική ευθεία.

(a) Υποθέστε ίσες a-priori πιθανότητες για κάθε κλάση.

(b) Υποθέστε a-priori πιθανότητες 2/3 για την κλάση A και 1/3 για την κλάση B.

5. Ταξινομητής Βάσει LMS (Linear Mean Square)

(a) Αρχικοποιήστε τα βάρη $\mathbf{w}$ και την πόλωση b σε μικρές τυχαίες τιμές (0.1). Χρησιμοποιήστε τον επαναληπτικό αλγόριθμο LMS για να ενημερώσετε τα βάρη με βάση την ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος. Εκτελέστε 3 επαναλήψεις και σχολιάστε τη συμπεριφορά του σφάλματος. Ποιά είναι η τελική τιμή των βαρών;

---

2 HMM

Έστω ένα κρυφό μοντέλο Markov με τρεις καταστάσεις 1, 2, 3 και δύο τύπους παρατηρήσεων H και T. Δίνεται ο παρακάτω πίνακας μεταβάσεων, $a_{ij} = p(q_t = j|q_{t-1} = i)$:

$$\begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.2 & 0.7 & 0.1 \\ 0.2 & 0.1 & 0.7 \end{pmatrix}$$

Δίνονται επίσης οι ακόλουθες πιθανότητες των παρατηρήσεων

$P(O\|q)$			
O / q	1	2	3
H	0.5	0.7	0.25
T	0.5	0.3	0.75

Οι a-priori πιθανότητες είναι ίσες με $\pi_1 = p(q_0 = 1) = \pi_2 = p(q_0 = 2) = \pi_3 = p(q_0 = 3) = \frac{1}{3}$.
α) Υπολογίστε την πιθανότητα $P(O_0 = H, O_1 = H, O_2 = T)$.
β) Υπολογίστε την πιο πιθανή σειρά καταστάσεων δεδομένης της σειράς παρατηρήσεων ($O_0 = H$, $O_1 = H$, $O_2 = T$), χρησιμοποιώντας τον αλγόριθμο Viterbi.
γ) Θέλουμε να χρησιμοποιήσουμε την ακόλουθη σειρά παρατηρήσεων $HHTTTHHTTTTTHHHH$ για να εκπαιδεύσουμε το κρυφό μοντέλο Markov. Αντί για τον πολύπλοκο αλγόριθμο forward-backward, θα ασχοληθούμε τον εξής ψευδο-EM αλγόριθμο που αναφέρεται συχνά ως εκπαίδευση Viterbi (Viterbi-training).

Expectation step: Αποκωδικοποίηση χρησιμοποιώντας τον αλγόριθμο Viterbi.
Maximization step: Μεγιστοποίηση συνολικής πιθανότητας καταστάσεων π

Δηλαδή, ο αλγόριθμος εκπαίδευσης Viterbi βρίσκει την πιο πιθανή σειρά καταστάσεων δεδομένης της σειράς παρατηρήσεων, και στη συνέχεια μεγιστοποιεί τη συνολική πιθανότητα της σειράς καταστάσεων $q_0q_1q_2$ που μόλις υπολογίσαμε και παρατηρήσεων $O_0O_1O_2$ που δινόνται. Το δεύτερο βήμα του αλγορίθμου είναι η συνήθης εκπαίδευση με μεγιστοποίηση πιθανότητας Maximum-Likelihood training που εφαρμόζεται σε κάθε Μπεϋσιοανό δίκτυο με όλες τις παραμέτρους (q, O) παρατηρήσεις.
δ) Ποιες είναι οι κύριες διαφορές μεταξύ του forward-backward και Viterbi-training και ποιος αλγόριθμος αναμένεται να έχει καλύτερα αποτελέσματα;

---

3 Backpropagation

Υποθέστε ότι έχουμε το ακόλουθο νευρωνικό δίκτυο. Οι κόμβοι που βρίσκονται μέσα σε μόνο κύκλο υποδηλώνουν μεταβλητές (για παράδειγμα η x_1 είναι μια μεταβλητή εισόδου, h_1 είναι μια ενδιάμεση μεταβλητή, και y είναι μια μεταβλητή εξόδου). Οι κόμβοι που βρίσκονται μέσα σε διπλό κύκλο υποδηλώνουν συναρτήσεις (για παράδειγμα το $\sum$ υπολογίζει το άθροισμα των εισόδων του και η σ αντιπροσωπεύει τη συνάρτηση logistic $\sigma(x) = \frac{1}{1+e^{-x}}$).

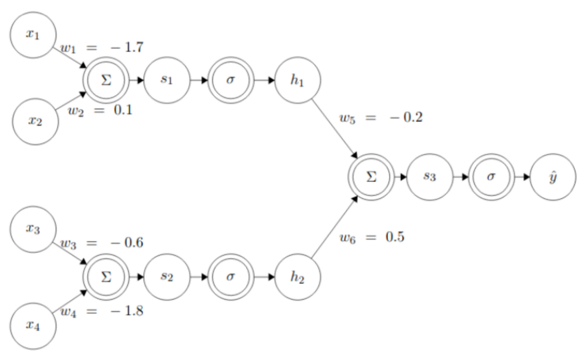

Figure 1: Σχηματική απεικόνιση του δικτύου

Θεωρήστε ότι η συνάρτηση για το L2 loss δίνεται από τη σχέση $L(y, \hat{y}) = ||y - \hat{y}||_2^2$. Επίσης, μας δίνονται τα δεδομένα ένα δείγματος $(x_1, x_2, x_3, x_4) = (-0.7, 1.2, 1.1, -2)$ με τιμή για το πραγματικό του label ίση με 0.5. Χρησιμοποιήστε τον αλγόριθμο backpropagation για να υπολογίσετε τη μερική παράγωγο $\frac{\partial L}{\partial w_1}$.
Σημείωση: Το gradient για μια συνάρτηση L2 loss είναι ίσο με $2||y - \hat{y}||$.

---

4 Expectation-Maximization

Θεωρήστε τα δεδομένα $D = \{(1, 2), (4, 5), (2, *)\}$ (το $*$ σημαίνει μια άγνωστη τιμή χαρακτηριστικού), τα οποία δειγματοληπτούνται από μια διδιάστατη διανομή $p(x_1, x_2) = p(x_1)p(x_2)$, με

$$p(x_1) \sim \begin{cases} \frac{1}{\theta_1} e^{-\theta_1 x_1} & \text{αν } x_1 \geq 0 \\ 0 & \text{αλλιώς} \end{cases}$$

$$p(x_2) \sim U(0, \theta_2) = \begin{cases} \frac{1}{\theta_2} & \text{αν } 0 \leq x_2 \leq \theta_2 \\ 0 & \text{αλλιώς} \end{cases}$$

a) Ξεκινώντας από μια αρχική εκτίμηση $\theta^0 = (2,3)$ υπολογίστε αναλυτικά το $Q(\theta, \theta^0)$, που αποτελεί το βήμα E του αλγορίθμου EM.

β) Βρείτε το θ που μεγιστοποιεί το $Q(\theta, \theta^0)$, που αποτελεί το βήμα M του αλγορίθμου EM.

γ) Σχεδιάστε ένα γράφημα που να δείχνει την κατανομή $p(x_1, x_2)$ πριν και μετά την εκτίμηση των παραμέτρων.

---

5 Bayes Network

Το Bayesian δίκτυο του Πνευμονολογικού Κέντρου (βλέπε το ακόλουθο σχήμα) αφορά τη διάγνωση της πνευμονικής νόσου (φυματίωση, καρκίνο του πνεύμονα ή και τα δύο, ή κανένα από τα δύο). Σε αυτό το μοντέλο υποθέτουμε ότι μια επίσκεψη στο μέρος Ω αυξάνει την πιθανότητα της φυματίωσης.

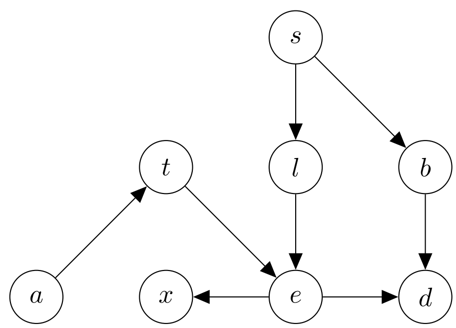

Figure 2: BN model of chest clinic

- x – Θετική ακτινογραφία X
- d – Δύσπνοια (Δυσκολία στην αναπνοή)
- e – Είτε Φυματίωση ή Καρκίνος του πνεύμονα
- t – Φυματίωση
- l – Καρκίνος του πνεύμονα
- b – Βρογχίτιδα
- a – Επίσκεψη στην περιοχή Ω
- s – Καπνιστής

(a) Παραγοντοποιήστε την κοινή κατανομή με βάση αυτό το μοντέλο BN.

(b) Επικυρώσετε ή απορρίψτε τις ακόλουθες δηλώσεις:

- καρκίνος του πνεύμονα $\perp\!\!\!\perp$ βρογχίτιδα | κάπνισμα
- επίσκεψη στο μέρος Ω $\perp\!\!\!\perp$ κάπνισμα | καρκίνος του πνεύμονα

(c) Υποθέστε ότι όλοι οι κόμβοι είναι δυαδικοί, δηλαδή, Αληθής ή Ψευδής. Οι υπό όρους πιθανότητες είναι οι εξής (όπου παραλείπουμε το συμβολισμό θεωρούμε τη δυαδική μεταβλητή ίση με True):

Έστω $P(a=T)=0.01$, $P(s=T)=0.5$, $P(t|a=T)=0.05$, $P(t|a=F)=0.01$, $P(l|s=T)=0.1$, $P(l|s=F)=0.01$, $P(b|s=T)=0.6$, $P(b|s=F)=0.3$, $P(x|e=T)=0.98$, $P(x|e=F)=0.05$, $P(d|e=T,b=T)=0.9$, $P(d|e=T,b=F)=0.7$, $P(d|e=F,b=T)=0.8$, $P(d|e=F,b=F)=0.1$, και $P(e=T|t,l)=0$ μόνο αν τόσο τα t όσο και l είναι 'Ψ', και 1 διαφορετικά.
Υπολογίστε τις τιμές για $P(x=T)$, $P(x=T|t=F)$, $P(t=T|x=F)$ και $P(t=T|x=F,d=T)$ με την χρήση του αλγόριθμου product-sum πάνω στον moralized/triangulated γράφο.

---

6 RNNs

Σε αυτή την άσκηση, θα εξερευνήσουμε ένα απλό αναδρομικό νευρωνικό δίκτυο (RNN) το οποίο μετρά τον αριθμό των 1 σε ένα δυαδικό ρεύμα εισόδου. Η κατάσταση σε κάθε χρονικό βήμα k δίνεται από τη σχέση αναδρομής:

$$s_k = s_{k-1} \cdot w_{rec} + x_k \cdot w_x$$

Σετ Δεδομένων Τα δεδομένα εισόδου X αποτελούνται από δυαδικές ακολουθίες μεταβλητού μήκους (μέγιστο μήκος ακολουθιας T), με την έξοδο t να αντιπροσωπεύει τον αριθμό (πλήθος) των 1 σε κάθε ακολουθία. Ουσιαστικά μετράμε τις μονάδες στην ακολουθία εισόδου.

Συνάρτηση Κόστους Υλοποιήστε την εμπρόσθια διάδοση (forward pass) για να υπολογίσετε τις καταστάσεις του δικτύου σε κάθε χρονικό βήμα k και χρησιμοποιήστε τη συνάρτηση απώλειας Mean Squared Error (MSE):

$$L(y,t) = (t-y)^2$$

Σημειώστε ότι θεωρούμε πως η συνάρτηση κόστους εφαρμόζεται μόνο στο τελευταίο timestep της ακολουθίας.

a) Υπολογίστε αναλυτικά (με βάση το MSE κόστος) τους κανόνες ανανέωσης των βαρών w_{rec} και w_x. Θεωρήστε μια ακολουθία που αποτελείται από T timesteps και το σφάλμα εφαρμόζεται στο τελικό timestep T.

b) Εκτελέστε 3 βήματα ανανέωσης των βαρών χειροκίνητα με αρχικά βάρη $w_{rec}=1$ και $w_x=0$, για την ακολουθία εισόδου [1,0,1,0,1,0].

c) Προτείνετε ένα διαφορετικό σετ βαρών αρχικοποίησης που θα μπορούσε να βελτιώσει την απόδοση του μοντέλου. Εκτελέστε ένα βήμα ανανέωσης με αυτό το νέο σετ βαρών. Σχολιάστε το αποτέλεσμα.

d) Υλοποιήστε το μοντέλο RNN της άσκησης και εκτελέστε την εκπαίδευση χρησιμοποιώντας τον κανόνα ανανέωσης που υπολογίσαμε στο βήμα α). Ξεκινώντας από 2-3 τυχαίες αρχικοποιήσεις βαρών, αναφέρετε τα αποτελέσματα και σχολιάστε.

e) **Bonus-1**: Με βάση την περιγραφή του συνδέσμου περιγράψτε τις διαφορές σε σχέση με την υλοποίηση της άσκησης ως προς τη συνάρτηση κόστους και τον τρόπου υπλογισμού των παραγώγων. Τί θα πρέπει να αλλάξετε στην υλοποίησή σας ώστε να πετύχετε BPTT όπως περιγράφεται στον υπερσύνδεσμο;

f) **Bonus-2**: Υλοποιήστε τις αλλαγές και τρέξτε κάποιες επαναλήψεις. Σχολιάστε τα αποτελέσματα.

ΒΙΒΛΙΟΓΡΑΦΙΑ:

[1] Γ. Καραγιάννης και Γ. Σταϊνχάουερ, *Αναγνώριση Προτύπων και Μάθηση Μηχανών*, ΕΜΠ, 2001.

[2] R. O. Duda, P.E. Hart and D.G. Stork, *Pattern Classification*, Wiley, 2001.

[3] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.

[4] S. Theodoridis and K. Koutroumbas, Pattern Recognition, 4th Edition Academic Pres, Elsevier, 2009. *Ελληνική μετάφραση*: απόδοση-επιμέλεια-πρόλογος ελληνικής έκδοσης Α. Πικράκης, Κ. Κουτρουμπάς, Θ. Γιαννακόπουλος, Επιστημονικές Εκδόσεις Π.Χ. Πασχαλίδης-Broken Hill Publishers LTD, 2012.

[5] S. Theodoridis, Machine Learning: a Bayesian and Optimization Perspective. Academic Press, 2015.