



# ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ & ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

Τομέας Επικοινωνιών, Ηλεκτρονικής & Συστημάτων Πληροφορικής

Εργαστήριο Διαχείρισης & Βέλτιστου Σχεδιασμού Δικτύων - NETMODE

Ηρώων Πολυτεχνείου 9, Ζωγράφου, 157 80 Αθήνα, Τηλ: 210-772.2503, Fax: 210-772.1452

e-mail: maglaris@netmode.ntua.gr, URL: <https://www.netmode.ntua.gr>

Γραπτή Εξέταση στο Μεταπτυχιακό Μάθημα  
"ΣΤΟΧΑΣΤΙΚΕΣ ΔΙΕΡΓΑΣΙΕΣ & ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ  
ΣΤΗ ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ"  
Εαρινό Εξάμηνο

02.07.2024

Διδάσκων: Β. Μάγκλαρης

Παρακαλώ απαντήστε σε όλες τις ερωτήσεις. Διάρκεια 2.5 ώρες

Στον τελικό βαθμό προσμετρώνται οι εργαστηριακές ασκήσεις και η προφορική εξέταση (40%)

ΚΑΛΗ ΕΠΙΤΥΧΙΑ!

## Θέμα 1 (5%)

Να αναφέρετε σε τι είδους μάθηση βασίζονται οι παρακάτω αλγόριθμοι:

- (α) Self-Organizing Map (SOM)
- (β) Temporal Difference Learning
- (γ) Radial Basis Function (RBF) Neural Network

## Θέμα 2 (7%)

(α) Να εξηγήσετε συνοπτικά πώς διαφέρουν οι αλγόριθμοι Hill Climbing και Simulated Annealing, που προσπαθούν να εντοπίσουν το ολικό ελάχιστο (global minimum) μίας συνάρτησης.

(β) Σε κάποιο βήμα του αλγορίθμου Simulated Annealing εξετάζετε τη μετάβαση του συστήματος από τη θέση  $x = e \approx 2.71$  στη θέση  $x = 1$ . Οι τιμές της ενέργειας του συστήματος δίνονται από τη συνάρτηση:  $E(x) = 10 \ln x - (\ln x)^2 + 1$ . Η θερμοκρασία του συστήματος είναι  $T = 30$ . Για να αποφασίσετε αν θα πραγματοποιηθεί το βήμα αυτό, παράγετε έναν τυχαίο αριθμό  $a$  από την ομοιόμορφη κατανομή  $[0, 1]$ . Για ποιες τιμές του  $a$  θα εκτελεστεί το βήμα;

πρόβλημα  
μετιστοποίηση

## Θέμα 3 (10%)

Να απαντήσετε συνοπτικά στις παρακάτω ερωτήσεις:



(α) Ο αλγόριθμος Metropolis-Hastings δημιουργεί δείγμα τυχαίων μεταβλητών από μία κατανομή πιθανοτήτων για την οποία υπάρχει αδυναμία υπολογισμού της σταθεράς κανονικοποίησης (partition function). Με ποιον τρόπο καταφέρνει ο Metropolis-Hastings να δώσει αποτελέσματα παρά το γεγονός ότι αγνοεί τη σταθερά κανονικοποίησης; Τι περιορισμό έχει το δείγμα τυχαίων μεταβλητών που δημιουργείται από τον αλγόριθμο;

(β) Να εξηγήσετε ποιοι από τους παρακάτω αλγορίθμους είναι parametric και ποιοι non-parametric: (i) K-Nearest Neighbors (KNN), (ii) Linear Regression, (iii) Naive Bayes

(γ) Ποια είναι η βασική διαφορά ανάμεσα στη μηχανή Boltzmann και στη Restricted Boltzmann Machine (RBM); Τα RBM's είναι διακριτικά (discriminative) ή παραγωγικά (generative) μοντέλα μηχανικής μάθησης;

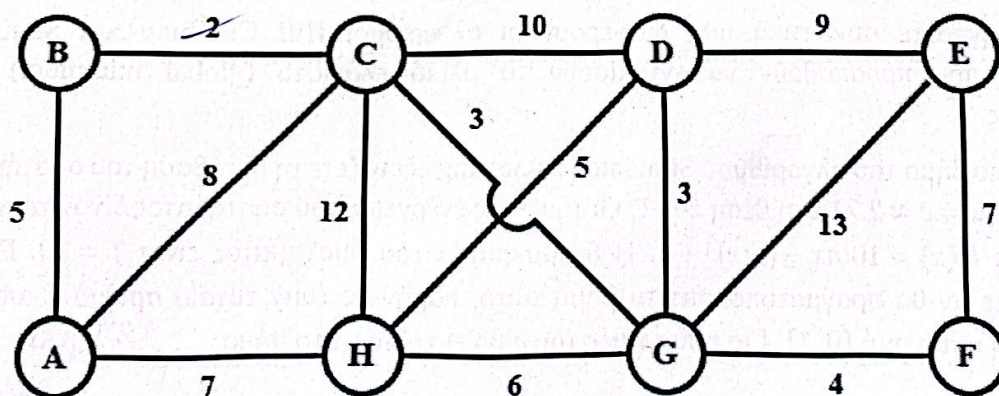
(δ) Ποια μοντέλα νευρωνικών δικτύων μπορούν να χρησιμοποιηθούν για τα επιμέρους επίπεδα (layers) ενός Deep Belief Network;

(ε) Πώς μπορεί να χρησιμοποιηθεί ένα Self-Organizing Map (SOM) σε εφαρμογές συμπίεσης εικόνων;

#### Θέμα 4 (10%)

(α) Ποια είναι η διαφορά ανάμεσα στους αλγορίθμους Value και Policy Iteration;

(β) Να εκτελέσετε τον αλγόριθμο Bellman-Ford για να βρείτε τις συντομότερες διαδρομές από τον κόμβο A προς τους υπόλοιπους κόμβους του δικτύου, όπως φαίνεται στον ακόλουθο γράφο. Οι ζεύξεις του δικτύου είναι διπλής κατεύθυνσης (full duplex) και έχουν συμμετρικά κόστη. Κατά την εκτέλεση του αλγορίθμου Bellman-Ford, οι κόμβοι του γράφου θα πρέπει να ανανεωθούν με την εξής σειρά: B, C, D, E, F, G, H.



#### Θέμα 5 (20%)

(α) Ποια είναι η παραδοχή που κάνει ο αλγόριθμος Naive Bayes Classifier και ποια πλεονεκτήματα προσφέρει;



(β) Βασισμένοι στον παρακάτω πίνακα, να εκπαιδεύσετε έναν Naive Bayes Classifier που θα αποφασίζει αν ένα σύνολο δικτυακών πακέτων με κοινά πεδία επικεφαλίδας (flows) σχετίζεται με κάποια κακόβουλη (ΚΑΚ) ή καλόβουλη (ΚΑΛ) δραστηριότητα. Το training set περιλαμβάνει 10000 flows. Τα 2 features που χρησιμοποιεί ο ταξινομητής είναι (i) ο συνολικός αριθμός πακέτων στο flow (μικρότερος/μεγαλύτερος από 3000 πακέτα, δηλαδή ΛΙΓΑ/ΠΟΛΛΑ) και (ii) ο μέσος χρόνος που μεσολαβεί ανάμεσα σε δύο διαδοχικά πακέτα (μικρότερος/μεγαλύτερος από 1 msec, δηλαδή ΜΙΚΡΟΣ/ΜΕΓΑΛΟΣ). Δίνεται ένα flow που αποτελείται από 1000 πακέτα με μέσο χρόνο διαφοράς δύο πακέτων ίσο με 0.5 msec. Ποια είναι η εκτίμηση του μοντέλου για αυτό το flow;

Αριθμός Πακέτων	Χρόνος	Καλόβουλα Flows (ΚΑΛ)	Κακόβουλα Flows (ΚΑΚ)
ΛΙΓΑ	ΜΙΚΡΟΣ	1400	600
ΠΟΛΛΑ	ΜΙΚΡΟΣ	400	2000
ΛΙΓΑ	ΜΕΓΑΛΟΣ	1000	1000
ΠΟΛΛΑ	ΜΕΓΑΛΟΣ	3100	500

Για την εκπαίδευση του αλγορίθμου να εκτιμήσετε τις prior και posterior πιθανότητες που είναι απαραίτητες για την ταξινόμηση του flow.

(γ) Σε περίπτωση εκπαίδευσης ενός Decision Tree για την επίλυση της παραπάνω ταξινόμησης, εξηγήστε ποιο feature (Αριθμός Πακέτων, Χρόνος) θα επιλέγατε ως ρίζα του Decision Tree. Να χρησιμοποιήσετε το Gini Index.

#### Θέμα 6 (8%)

Δίνεται ο πίνακας πιθανοτήτων μετάβασης για μία αλυσίδα Markov. Να συμπληρώσετε τα κενά του πίνακα, να προσδιορίσετε την επαναληπτική κλάση της αλυσίδας και να βρείτε την αναλλοίωτη κατανομή των επαναληπτικών καταστάσεων.

$$\begin{pmatrix}
 ? & 1/4 & 0 & 0 & 0 & 1/2 \\
 ? & 1/2 & 1/2 & ? & ? & 0 \\
 0 & 1/3 & 0 & 0 & 0 & ? \\
 0 & 0 & 1/6 & ? & 1/3 & 0 \\
 1/7 & 2/7 & 0 & ? & 0 & 0 \\
 0 & 1/4 & ? & 0 & 0 & 0
 \end{pmatrix}$$