



ΥΠΟΛΟΓΙΣΤΙΚΗ ΝΟΗΜΟΣΥΝΗ

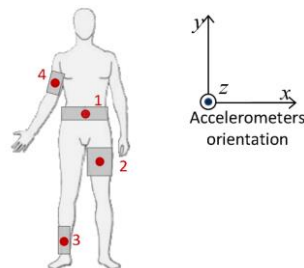
Διδάσκοντες: Σ. Λυκοθανάσης, Δ. Κουτσομητρόπουλος

Ακαδημαϊκό Έτος 2022-2023

Εργαστηριακή Άσκηση Μέρος Α'

Α. Αναγνώριση ανθρώπινης δραστηριότητας με Χρήση Νευρωνικών Δικτύων

Η αναγνώριση ανθρώπινης δραστηριότητας (Human Activity Recognition) αποτελεί μια διεργασία ερμηνείας της ανθρώπινης κίνησης από τα υπολογιστικά συστήματα. Οι δραστηριότητες καταγράφονται με τη βοήθεια αισθητήρων, ενώ τα δεδομένα που προκύπτουν αναλύονται-αξιοποιούνται από διάφορους τομείς όπως υγεία, αθλητισμός, ψυχαγωγία, κ.α. Μια χαρακτηριστική περίπτωση αποτελεί και η προσπάθεια ταξινόμησης στάσεων-κινήσεων του σώματος σε ένα προκαθορισμένο πλήθος κλάσεων. Συγκεκριμένα, στην παρούσα εργασία σας δίνονται δεδομένα από τέσσερις αισθητήρες που τοποθετήθηκαν σε τέσσερις χρήστες για συνολικά 8 ώρες, και σας ζητείται υλοποίηση που να ταξινομεί τα δεδομένα αυτά σε ακριβώς μια από τις πέντε διαθέσιμες κλάσεις. Στην ακόλουθη εικόνα εμφανίζονται οι θέσεις των τεσσάρων αισθητήρων (φορητά επιταχυνσιόμετρα) πάνω στο ανθρώπινο σώμα.



Ειδικότερα, πρέπει να εξετάσετε τη χρήση ενός πολυεπίπεδου ΤΝΔ για την πρόβλεψη των 5 κλάσεων (*sitting-down*, *standing-up*, *standing*, *walking*, και *sitting*) που αντιστοιχούν σε κάθε δείγμα. Για τον σκοπό αυτό θα αξιοποιηθεί το σύνολο δεδομένων *PUC-Rio*¹ που συγκεντρώνει 165.633 δείγματα για εκπαίδευση και έλεγχο. Οι ακέραιες τιμές από τους αισθητήρες μαζί με τα βασικά στοιχεία των χρηστών, από τους οποίους έγινε η λήψη, αλλά και το είδος της δραστηριότητας (κλάση) περιέχονται στο αρχείο *dataset-HAR-PUC-Rio.csv*.

Αναλυτικά, κάθε εγγραφή-γραμμή περιλαμβάνει τα εξής στοιχεία: χρήστης, φύλο, ηλικία, ύψος, βάρος, δείκτης μάζας σώματος, (x,y,z) από αισθητήρα μέσης, (x,y,z) από αισθητήρα αριστερού μηρού, (x,y,z) από αισθητήρα δεξιού αστράγαλου, (x,y,z) από αισθητήρα πάνω μέρους χεριού και κλάση. Οπότε π.χ. ένα δείγμα έχει ως εξής :

wallace;Man;31;1,71;83;28,4;-24;92;-118;1;92;-128;-12;166;-87;-212;-74;-173;walking

Για την υλοποίηση των αλγορίθμων μπορείτε να χρησιμοποιήσετε οποιοδήποτε περιβάλλον, βιβλιοθήκη ή γλώσσα προγραμματισμού κρίνετε σκόπιμο. Ενδεικτικά αναφέρονται: *MatLab*, *WEKA*, *Azure ML Studio*, *Google Colaboratory*, *TensorFlow*, *Keras*, *SciKit-Learn*.

¹ <https://eclass.upatras.gr/modules/document/file.php/CEID1060/dataset-HAR-PUC-Rio.csv>

[UCI Machine Learning Repository: Wearable Computing: Classification of Body Postures and Movements \(PUC-Rio\) Data Set](#)

Το ζητούμενο στην εργασία αυτή είναι να κατασκευαστεί και να εκπαιδευτεί ένα ΤΝΔ που να ταξινομεί εισόδους (δραστηριότητες) σε 5 διαφορετικές κλάσεις.

A1. Προεπεξεργασία και Προετοιμασία δεδομένων [20 μονάδες]

Προσοχή: Ό,τι μετασχηματισμοί εφαρμοστούν στα δεδομένα του συνόλου εκπαίδευσης, οι ίδιοι θα πρέπει να εφαρμοστούν και στα δεδομένα του συνόλου ελέγχου ή εναλλακτικά να αντιστραφούν πρώτου μετρηθούν οι μετρικές αξιολόγησης παρακάτω.

α) *Κωδικοποίηση και προεπεξεργασία δεδομένων*: Στο σύνολο δεδομένων υπάρχουν κυρίως αριθμητικές, αλλά και κατηγορικές τιμές (φύλο, στάση σώματος). Οι κατηγορικές τιμές, προκειμένου να μπορέσουν να χρησιμοποιηθούν στην εκπαίδευση ενός ΤΝΔ θα πρέπει να αντιστοιχηθούν σε αριθμητικές. Για παράδειγμα, οι 5 κλάσεις (*sitting-down*, *standing-up*, *standing*, *walking*, *sitting*) μπορούν να αντικατασταθούν με τις τιμές (1, 2, 3, 4, 5) σε όλο το σύνολο δεδομένων. Οι τιμές από τους αισθητήρες, όπως αναφέρθηκε, είναι ακέραιες και κινούνται στο διάστημα [-617,533]. Το εύρος τιμών των δεδομένων μπορεί να διαφέρει σημαντικά ανά χαρακτηριστικό. Για αυτό τον λόγο, υπάρχει κίνδυνος να υπερεκτιμηθεί η συνεισφορά κάποιου χαρακτηριστικού έναντι άλλων ή να υπάρξει πόλωση στις τιμές κάποιων χαρακτηριστικών ή στις τιμές των βαρών που θα προκύψουν. Με δεδομένη την συγκεκριμένη αποτύπωση και την πιθανή ανάγκη προσαρμογής των τιμών αυτών σε διαφορετική κλίμακα (εξάλειψη ενδεχόμενης πόλωσης), παρουσιάζονται οι εξής τρεις μέθοδοι:

- *Κεντράρισμα (Centering)*: Με την μέθοδο αυτή αφαιρούμε τον μέσο όρο των τιμών για κάθε χαρακτηριστικό από όλες τις τιμές που έχουν αποδοθεί.
- *Κανονικοποίηση (Normalization ή min-max scaling)*: Με την μέθοδο αυτή μεταφέρουμε το εύρος τιμών ενός χαρακτηριστικού σε νέα κλίμακα πχ [0,1].
- *Τυποποίηση (Standardization ή z-score)*: Με την μέθοδο αυτή παρέχουμε στο δείγμα ιδιότητες όπως μηδενική μέση τιμή και μοναδιαία διακύμανση (Gaussian).

Εξετάστε τη χρησιμότητα των ανωτέρω μεθόδων για το συγκεκριμένο πρόβλημα και εφαρμόστε τη/τις στα δεδομένα εκπαίδευσης, αν κρίνετε σκόπιμο. [15]

β) *Διασταυρούμενη Επικύρωση (cross-validation)*: Βεβαιωθείτε ότι έχετε διαχωρίσει τα δεδομένα σας σε σύνολα εκπαίδευσης και ελέγχου, ώστε να χρησιμοποιήσετε 5-fold CV για όλα τα πειράματα. Προσέξτε το κάθε fold να είναι *ισορροπημένο* (balanced) ως προς τον αριθμό των δειγμάτων κάθε κλάσης. [5]

A2. Επιλογή αρχιτεκτονικής [50 μονάδες]

Όσον αφορά την τοπολογία των ΤΝΔ για την εκπαίδευση τους με τον Αλγόριθμο Οπισθοδιάδοσης του Σφάλματος (back-propagation), θα χρησιμοποιήσετε ΤΝΔ με *ένα κρυφό επίπεδο* και θα πειραματιστείτε με τον αριθμό των κρυφών κόμβων. Για την εκπαίδευση του δικτύου χρησιμοποιήστε αρχικά ρυθμό μάθησης $\eta = 0.001$.

α) Η εκπαίδευση και αξιολόγηση των μοντέλων σας μπορεί να γίνει με χρήση *Cross-Entropy* (CE), *Μέσου Τετραγωνικού Σφάλματος* (MSE), καθώς και *ακρίβεια ταξινόμησης* (Accuracy)². Να εξηγήσετε με απλά λόγια ποια είναι η σημασία των παραπάνω μετρικών για το συγκεκριμένο πρόβλημα. Ποια είναι προτιμότερη για εκπαίδευση (loss); [5]

β) Πόσους νευρώνες θα χρειαστείτε στο επίπεδο εξόδου, δεδομένου του ζητούμενου της ταξινόμησης σε πολλαπλές κλάσεις (multiclass classification); [5]

² Πρόκειται για το ποσοστό των ορθών προβλέψεων από το δίκτυο. Αν \hat{y}_i είναι η επιθυμητή κλάση για το δείγμα i και y_i η κλάση που προβλέπει το δίκτυο, τότε $\text{accuracy}(y, \hat{y}) = \frac{1}{n_{\text{samples}}} \sum_{i=0}^{n_{\text{samples}}-1} 1(y_i = \hat{y}_i)$

γ) Να επιλέξετε κατάλληλη συνάρτηση ενεργοποίησης για τους κρυφούς κόμβους και να τεκμηριώσετε την επιλογή σας. [5]

δ) Ποια συνάρτηση ενεργοποίησης θα χρησιμοποιήσετε για το επίπεδο εξόδου; Σιγμοειδή, γραμμική, **Softmax** ή κάποια άλλη; [5]

ε) Πειραματιστείτε με 3 διαφορετικές τιμές για τον αριθμό των νευρώνων του κρυφού επιπέδου και συμπληρώστε τον παρακάτω πίνακα. Εμπειρικά ενδεδειγμένες τιμές για τον αριθμό των κρυφών κόμβων βρίσκονται στο διάστημα $[O, I+O]$ (I αριθμός εισόδων, O αριθμός εξόδων, H αριθμός κόμβων στο κρυφό επίπεδο). Να συμπεριλάβετε και τις γραφικές παραστάσεις σύγκλισης (Μ.Ο.) ανά κύκλο εκπαίδευσης. Διατυπώστε τα συμπεράσματά σας σχετικά με (i) τον αριθμό των κρυφών κόμβων, (ii) την επιλογή της συνάρτησης κόστους και (iii) την ταχύτητα σύγκλισης ως προς τις εποχές εκπαίδευσης. [15]

Αριθμός νευρώνων στο κρυφό επίπεδο	CE loss	MSE	Acc
$H_I = O$			
$H_I = (I+O)/2$			
$H_I = I+O$			

στ) Κριτήριο τερματισμού. Επιλέξτε και τεκμηριώστε κατάλληλο κριτήριο τερματισμού της εκπαίδευσης κάθε φορά (για κάθε fold). Μπορεί να χρησιμοποιηθεί η τεχνική του πρόωρου σταματήματος (early stopping); [5]

Προσοχή: σε όλα τα πειράματα θα χρησιμοποιήσετε 5-fold cross validation (5-fold CV).

A3. Μεταβολές στον ρυθμό εκπαίδευσης και σταθεράς ορμής [15 μονάδες]

Επιλέγοντας την τοπολογία που δίνει το καλύτερο αποτέλεσμα βάσει του προηγούμενου ερωτήματος, να πραγματοποιήσετε βελτιστοποίηση των υπερπαραμέτρων ρυθμού εκπαίδευσης η και σταθεράς ορμής m με χρήση CV και να συμπληρώσετε τον παρακάτω πίνακα. Να συμπεριλάβετε και τις γραφικές παραστάσεις σύγκλισης (Μ.Ο.) ως προς τους κύκλους εκπαίδευσης που θα χρειαστούν. Να τεκμηριώσετε θεωρητικά γιατί $m < 1$.

η	m	CE loss	MSE	Acc
0.001	0.2			
0.001	0.6			
0.05	0.6			
0.1	0.6			

Να διατυπώσετε σύντομα τα συμπεράσματα που προκύπτουν από τα 4 πειράματα.

A4. Ομαλοποίηση [15 μονάδες]

Μια μέθοδος για την αποφυγή υπερπροσαρμογής του δικτύου και βελτίωση της γενικευτικής του ικανότητας είναι η ομαλοποίηση του διανύσματος των βαρών (regularization). Να εξηγήσετε ποια μέθοδο ομαλοποίησης ($L1$ ή $L2$) είναι προτιμότερη για το συγκεκριμένο πρόβλημα. Στη συνέχεια να την εφαρμόσετε και να επανεκπαιδεύσετε το δίκτυό σας, όπως προέκυψε από το A3, αξιολογώντας διάφορες τιμές για τον συντελεστή r .

i) $r = 0.1$ ii) $r = 0.5$ iii) $r = 0.9$

Συμπληρώστε τον παρακάτω πίνακα για κάθε μία από τις παραπάνω περιπτώσεις με χρήση 5-fold CV. Να συμπεριλάβετε και τις γραφικές παραστάσεις σύγκλισης (Μ.Ο.) ανά κύκλο εκπαίδευσης.

Συντελεστής r	CE loss	MSE	Acc
0.1			
0.5			
0.9			

Διατυπώστε τα συμπεράσματά σας σχετικά με την επίδραση της μεθόδου στη γενικευτική ικανότητα του δικτύου.

A5. Βαθύ Νευρωνικό Δίκτυο [προαιρετικό ερώτημα - 10 μονάδες bonus]

Δοκιμάστε να προσθέσετε περισσότερα του ενός κρυφά επίπεδα στο δίκτυο (μέχρι 3). Πειραματιστείτε με τον αριθμό των κόμβων, όπως κάνατε στο A2. Περιγράψτε μια λογική για την στοίχιση των κρυφών επιπέδων (είναι καλό να έχουν τον ίδιο αριθμό κόμβων; Μειούμενο; Αυξανόμενο;). Να αναφέρετε CE, MSE και Acc για τα πειράματά σας με 5-fold CV και να διατυπώσετε τα συμπεράσματά σας σχετικά με την προσθήκη κρυφών επιπέδων.

Παραδοτέα

Η αναφορά που θα παραδώσετε θα πρέπει να περιέχει εκτενή σχολιασμό των πειραμάτων σας, καθώς και πλήρη καταγραφή των αποτελεσμάτων και των συμπερασμάτων σας, ανά υπο-ερώτημα. Επίσης, πρέπει να συμπεριλάβετε στην αρχή της αναφοράς σας ένα link προς τον κώδικα που έχετε χρησιμοποιήσει (σε κάποια file sharing υπηρεσία ή code repo).

Μην ξεχάσετε να συμπληρώσετε τα στοιχεία σας στην αρχή της 1^{ης} σελίδας.

Αξιολόγηση

Η απάντηση των ερωτημάτων Α και Β έχει βαρύτητα 20% στον τελικό βαθμό του μαθήματος (το σύνολο και των δύο μερών της εργασίας έχει βαρύτητα 40%). Ο βαθμός του Bonus (10%) προστίθεται στο παραπάνω ποσοστό 40%.

Παρατηρήσεις

1. Η αναφορά, σε ηλεκτρονική μορφή, πρέπει να αναρτηθεί στο e-class μέχρι τη Δευτέρα, 24/4/2023, στις 23:59.
2. Για οποιαδήποτε διευκρίνιση / ερώτηση μπορείτε να χρησιμοποιείτε το σχετικό forum στο eclass του μαθήματος.