
AIoT Course - Human Gesture Recognition Project

·ονοματεπώνυμο : Ανδρέας Καρεσίου

·AM : 1078717

·έτος : 4^ο

·ονοματεπώνυμο : Ανδρέας Φωτάκης

·AM : 1084674

·έτος : 4^ο

GYM MOVEMENTS RECOGNITION-CLASSIFICATION

Αρχικά, συλλέξαμε δεδομένα για τις ακόλουθες κινήσεις γυμναστικής: bicep curl, shoulder press, lateral raises, και row. Ενώσαμε τα αρχεία acc και gyr κάθε συλλογής σε ένα αρχείο που περιέχει και τους έξι άξονες. Στη συνέχεια, εφαρμόσαμε κανονικοποίηση (normalization) των τιμών σε κάθε άξονα και φιλτράρισμα χαμηλών συχνοτήτων (low pass filtering). Υπολογίσαμε την ενέργεια βάσει των έξι αξόνων για κάθε χρονική στιγμή. Έπειτα, κάναμε segmentation σε όλες τις κινήσεις κάθε συλλογής και αποθηκεύσαμε κάθε τμήμα σε ξεχωριστό αρχείο. Σε κάθε τμήμα εφαρμόσαμε sliding window με επικάλυψη 20% (70 length) και αποθηκεύσαμε τα παράθυρα τοπικά, προτού τα ανεβάσουμε στη MongoDB. Κάθε συλλογή περιείχε κινήσεις μόνο μιας κατηγορίας, γεγονός που διευκόλυνε τον διαχωρισμό. Για το segmentation, εκτός από την ενέργεια, χρησιμοποιήσαμε και τα εύρη τιμών στους άξονες για να εντοπίσουμε τις περιόδους ακινησίας σε κάθε χρονοσειρά(σε αρκετές περιπτώσεις κάποιες τιμές τις αλλάξαμε χειρωνακτικά). Στη συνέχεια, πήραμε τα windowed segments από τη βάση για κάθε κατηγορία και τα αποθηκεύσαμε σε έναν πίνακα. Επειδή γνωρίζαμε εκ των προτέρων τα indices για κάθε κίνηση και τον αριθμό των windowed segments κάθε κίνησης, ήταν εύκολο να αντιστοιχίσουμε τις τιμές για το classification:

```
y = np.ones(2543, dtype=int)
```

```
y[1:563]=0 # bicep curl
```

```
y[564:1283]=1 # shoulder press
```

```
y[1284:2017]=2 # lateral raises
```

```
y[2018:2543]=3 # row
```

Στη συνέχεια, τα δεδομένα διαχωρίζονται σε σύνολα εκπαίδευσης και ελέγχου. Εφαρμόζεται κλιμάκωση στα δεδομένα για να βελτιωθεί η απόδοση των αλγορίθμων μηχανικής μάθησης, χρησιμοποιώντας τον StandardScaler της βιβλιοθήκης sklearn. Η ανάλυση κύριων συνιστωσών (PCA) χρησιμοποιείται για να μειωθούν οι διαστάσεις των δεδομένων, διευκολύνοντας την εκπαίδευση και την αξιολόγηση των μοντέλων. Αρχικά, πραγματοποιείται PCA με μία συνιστώσα για να απλοποιηθούν τα δεδομένα και να κατανοηθεί καλύτερα η δομή τους. Στη συνέχεια, εφαρμόζεται PCA με τρεις συνιστώσες για οπτικοποίηση των δεδομένων σε τρεις διαστάσεις.

Για την αρχική εκπαίδευση και αξιολόγηση των δεδομένων χρησιμοποιούνται απλοί ταξινομητές, όπως ο SVM (Support Vector Machine) και το Random Forest. Οι ταξινομητές αυτοί εκπαιδεύονται με τα δεδομένα εκπαίδευσης και αξιολογούνται με τα δεδομένα ελέγχου. Οι μετρήσεις απόδοσης, όπως η ακρίβεια και οι ταξινομητικές αναφορές, εκτυπώνονται για να αξιολογηθεί η απόδοση των μοντέλων. Εφαρμόζεται επίσης βελτιστοποίηση με Grid Search για να βρεθούν οι καλύτεροι παράμετροι του SVM ταξινομητή. Αυτό γίνεται μέσω της GridSearchCV της sklearn, η οποία επιτρέπει την εξαντλητική αναζήτηση των καλύτερων συνδυασμών παραμέτρων για το μοντέλο.

Στη συνέχεια, δημιουργούνται και εκπαιδεύονται νευρωνικά δίκτυα για την ταξινόμηση των δεδομένων. Το πρώτο νευρωνικό δίκτυο που δημιουργείται είναι ένα convolutional neural network(CNN) Το μοντέλο εκπαιδεύεται για 100 epoches με batch size 64 και η απόδοσή του αξιολογείται μέσω accuracy and loss. Η εκπαίδευση και η αξιολόγηση των αποτελεσμάτων οπτικοποιούνται με γραφήματα. Τα γραφήματα αυτά δείχνουν την απώλεια και την ακρίβεια κατά την εκπαίδευση και την επικύρωση, επιτρέποντας την κατανόηση της απόδοσης του μοντέλου σε κάθε epoch. Επιπλέον, δημιουργείται και εκπαιδεύεται ένα απλό Recursive neural network(RNN). Το μοντέλο εκπαιδεύεται με τα ίδια δεδομένα εκπαίδευσης και αξιολογείται με τα ίδια δεδομένα ελέγχου όπως και το CNN. Η απόδοση του RNN συγκρίνεται με αυτή του CNN μέσω γραφημάτων που δείχνουν την απώλεια και την ακρίβεια κατά την εκπαίδευση και την επικύρωση.

Τέλος, εκτυπώνονται τα confusion tables για τα δεδομένα εκπαίδευσης και ελέγχου για να αξιολογηθεί η απόδοση των μοντέλων στην πρόβλεψη των κατηγοριών. Δείχνουν πόσα παραδείγματα κάθε κατηγορίας προβλέφθηκαν σωστά και πόσα λανθασμένα.

Παρατηρήσεις για τα νευρωνικά δίκτυα:

Το CNN δείχνει πιο σταθερές καμπύλες εκμάθησης, με τη μείωση της απώλειας και την αύξηση της ακρίβειας να είναι πιο ομαλές. Η εκπαίδευση του CNN φτάνει σε πολύ υψηλά επίπεδα ακρίβειας, ενώ η απώλεια πέφτει κοντά στο μηδέν. Από την άλλη πλευρά, το RNN παρουσιάζει περισσότερες διακυμάνσεις και αστάθεια κατά τη διάρκεια της εκπαίδευσης. Η απώλεια παρουσιάζει αιχμές, και η ακρίβεια επίσης μεταβάλλεται σημαντικά, υποδεικνύοντας πιθανή δυσκολία στην εκμάθηση συγκεκριμένων χαρακτηριστικών των δεδομένων.

Επειδή οι κινήσεις(παρά την σημαντική προσπάθεια μας να τις κάνουμε όσο πιο διαφορετικές μπορούσαμε(κάθε κατηγορίας)) είχαν αρκετά όμοια κομμάτια οπότε κατά την εκπαίδευση παρουσιάστηκαν προβλήματα overfitting τα οποία με trial and error μέθοδο τα μειώσαμε αρκετά εφαρμόζοντας dropout τεχνική