

Εθνικό Μετσόβιο Πολυτεχνείο

Σχολή Ηλεκτρολόγων Μηχανικών & Μηχανικών Υπολογιστών

Δ.Π.Μ.Σ. Επιστήμη Δεδομένων & Μηχανική Μάθηση



Εξαμηνιαία Εργασία στη Μηχανική Μάθηση

Γιώργος Βερνίκος Α.Μ.: 03400005

Νικηφόρος Μανδηλαράς Α.Μ.: 03400022

Χρήστος Σπυρόπουλος Α.Μ.: 03400035

Παραδοτέα 3/3/19

1 ΕΙΣΑΓΩΓΗ

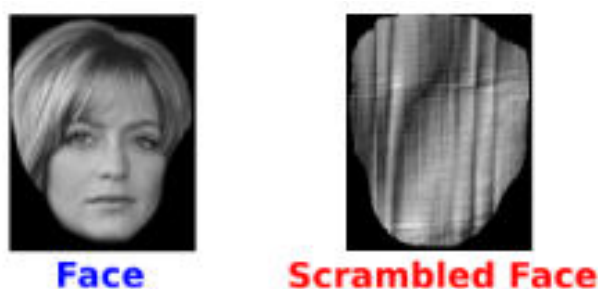
Ο διαγωνισμός τον οποίο επιλέξαμε για την εκπόνηση της εξαμηνιαίας εργασίας της ομάδας μας στα πλαίσια του μαθήματος «Μηχανική Μάθηση» του Δ.Π.Μ.Σ «Επιστήμη Δεδομένων και Μηχανική Μάθηση» ονομάζεται *DecMeg2014 – Decoding the Human Brain* και διεξήχθη μέσω της πλατφόρμας *Kaggle* το 2014. Όπως μαρτυρά το όνομά του, πρόκειται για ένα διαγωνισμό του οποίου στόχος είναι η διερεύνηση του τρόπου λειτουργίας τους ανθρωπίνου εγκεφάλου, μέσω ενός πειράματος.

Ο στόχος αυτός αποτελεί βασικό αντικείμενο της Νευροεπιστήμης, η οποία τα τελευταία χρόνια έχει συνδεθεί και με τον τομέα της τεχνητής νοημοσύνης δημιουργώντας έναν νέο κλάδο ανάμεσα στις βιολογικές επιστήμες και την μαθηματική μοντελοποίηση, αυτόν της Υπολογιστικής Νευροεπιστήμης (Computational Neuroscience). Στόχος αυτού του κλάδου είναι η χρήση μοντέλων για την κατανόηση της δομής, της λειτουργίας και της γνωστικής διαδικασίας του νευρικού μας συστήματος. Κάποιοι, μάλιστα, από τους αλγορίθμους που χρησιμοποιούνται ακόμα και τώρα στα Νευρωνικά Δίκτυα, είχαν αρχικά αυτό το σκοπό, εξ ου και το όνομά τους. Τις περισσότερες φορές για τον έλεγχο υποθέσεων ή τη θεμελίωση μοντέλων, επιστήμονες του κλάδου αυτού διεξάγουν πειράματα όπως αυτό από το οποίο προέρχονται τα δεδομένα του διαγωνισμού.

Το συγκεκριμένο πείραμα κάνει χρήση του μαγνητοεγκεφαλογράφηματος (MEG), μιας τεχνικής που προσπαθεί να απεικονίσει την εγκεφαλική δραστηριότητα ενός υποκειμένου, μετρώντας την ισχύ των μαγνητικών πεδίων που δημιουργούνται, ως συνέπεια των ηλεκτρικών ρευμάτων που διαρρέουν τους νευρώνες με τη βοήθεια πολύ ευαίσθητων αισθητήρων. Απώτερος στόχος αυτής της διαδικασίας είναι η σύνδεση της εγκεφαλικής δραστηριότητας του υποκειμένου με το αντίστοιχο ερέθισμα που την προκάλεσε, ως ένα πρώτο βήμα για τη χαρτογράφηση του εγκεφάλου και την αποκωδικοποίηση (decoding) του τρόπου λειτουργίας του. Η συγκεκριμένη μέθοδος χρησιμοποιείται κατά κόρον στην Αλληλεπίδραση Ανθρώπου-Μηχανής και πιο συγκεκριμένα στην κατηγορία προβλημάτων στην οποία ο άνθρωπος προσπαθεί μέσω της σκέψης του να χειριστεί μια μηχανή. Η προαναφερθείσα μέθοδος αλληλεπίδρασης βασίζεται στο γεγονός ότι θα υπάρξει κάποια αλλαγή στην ενεργοποίηση ορισμένων νευρώνων κατά την εκτέλεση μιας κίνησης αλλά και κατά την σκέψη της εκτέλεσής της. Αυτή η αλλαγή (Event-Related Desynchronization) εκφράζεται σε όρους επεξεργασίας σημάτων ως μια μετάβαση από μια κατάσταση ηρεμίας (idle state) σε μια κατάσταση υψηλής ενέργειας. Έτσι, στα περισσότερα συστήματα στα οποία ένας άνθρωπος ελέγχει μια μηχανή μόνο με τη σκέψη του γίνεται αρχικά μια προσαρμογή στα εγκεφαλικά σήματα του

υποκειμένου και ύστερα με βάση την εγκεφαλική του δραστηριότητα μπορεί να καθοδηγήσει τη λειτουργία της μηχανής.

Το πείραμα του διαγωνισμού αφορά την ανίχνευση και αναγνώριση της εγκεφαλικής δραστηριότητας ενός υποκειμένου κατά την παρουσίαση εικόνων σε αυτό. Πιο συγκεκριμένα στο υποκείμενο παρουσιάζονται εικόνες οι οποίες ανήκουν σε μία από τις δύο κατηγορίες: είτε αποτελούν εικόνες προσώπων ή αποτελούν εικόνες προσώπων στις οποίες τα εικονοστοιχεία (pixels) έχουν ανακατευτεί με τέτοιο τρόπο, έτσι ώστε να μη φαίνεται πλέον ότι είναι εικόνα προσώπου όπως φαίνεται και στο παρακάτω σχήμα.



Εικόνα 1. Παράδειγμα των διαφορετικών εικόνων που παρουσιάζονται στα υποκείμενα

Στόχος, του διαγωνισμού είναι η χρήση των μετρήσεων των 306 αισθητήρων που ήταν τοποθετημένοι σε κάθε υποκείμενο για την αναγνώριση του ερεθίσματος που προκάλεσε την εκάστοτε εγκεφαλική διέγερση, δηλαδή η πρόβλεψη για το αν το αντικείμενο έβλεπε μια κανονική ή μια αλλοιωμένη εικόνα. Κάθε διέγερση και αντίστοιχη καταγραφή του μαγνητοεγκεφαλογραφήματος του υποκειμένου αποτελεί μία δοκιμή και για κάθε υποκείμενο έγιναν αρκετές δοκιμές. Κάποια υποκείμενα εμφανίζονται στο σετ εκπαίδευσης (training set) ενώ κάποια άλλα εμφανίζονται στο σετ εξέτασης (test set) για να αξιολογηθεί η ικανότητα γενίκευσης του εκάστοτε μοντέλου.

2 ΣΧΕΤΙΚΕΣ ΜΕΘΟΔΟΙ

Η διαδικασία που συνηθίζεται στη βιβλιογραφία για το πρόβλημα της ανάλυσης μαγνητοεγκεφαλογραφήματος αλλά και ηλεκτροεγκεφαλογραφήματος (EEG) βασίζεται στη χρήση μεθόδων Αναγνώρισης Σημάτων για την προεπεξεργασία των σημάτων και Μηχανικής Μάθησης για τη δημιουργία ταξινομητών για τα εξαγόμενα χαρακτηριστικά. Βασικά προβλήματα της επεξεργασίας αυτών των σημάτων είναι το ότι τα παραγόμενα σήματα διαφέρουν από υποκείμενο σε υποκείμενο, η εγκεφαλική λειτουργία παράγει έτσι κι αλλιώς κάποια ηλεκτρικά σήματα οπότε είναι δύσκολο να απομονωθεί το ερέθισμα καθώς και το γεγονός ότι για την ταξινόμηση ενός ερεθίσματος είναι απαραίτητος ο περιορισμός της περιοχής του εγκεφάλου που το αποκωδικοποιεί. Έτσι, η διαδικασία της

επεξεργασίας του σήματος έγκειται στην προβολή των σημάτων σε έναν υποχώρο μέσω κάποιων χωρικών φίλτρων για την απομόνωση του ερεθίσματος στην εκάστοτε περιοχή. Αυτό συνήθως γίνεται με τη χρήση του αλγορίθμου Common Spatial Patterns (CSP) (Benjamin Blankertz, 2008) ο οποίος αναλύει πολυκάναλα σήματα (σήματα που προέρχονται από πολλούς αισθητήρες) με βάση δοκιμές για δύο κλάσεις. Ως μέθοδος, μεγιστοποιεί τη διασπορά του χωρικά φιλτραρισμένου σήματος για τη μία κλάση, ελαχιστοποιώντας την παράλληλα για την άλλη κλάση. Έχουν εμφανιστεί επεκτάσεις και βελτιώσεις του παραπάνω αλγορίθμου, όπως το φιλτράρισμα στη συχνότητα των σημάτων ανάλογα με το εκάστοτε υποκείμενο (Kai Keng Ang C. G., 2008) (Kai Keng Ang, 2012). Τέλος, με την άνθηση των Νευρωνικών Δικτύων τα τελευταία χρόνια, έχει γίνει χρήση τους και σε αυτό το πρόβλημα, με τη χρήση κυρίως συνελκτικών δικτύων σε συνδυασμό με τις παραδοσιακές μεθόδους της Επεξεργασίας Σημάτων (Siavash Sakhavi, 2018).

3 Η ΜΕΘΟΔΟΣ ΜΑΣ

Στα πλαίσια της παρούσας εργασίας και με τη ραγδαία ανάπτυξη, πλέον, των Νευρωνικών Δικτύων προτείνουμε μια διαφορετική προσέγγιση για την επίλυση του προβλήματος. Σε αντίθεση με τις παραδοσιακές μεθόδους αλλά και κάποιες πρόσφατες υλοποιήσεις, δε κάνουμε χρήση της γνώσης της Επεξεργασίας Σήματος και αντ' αυτού προσπαθούμε να εφαρμόσουμε μεθόδους Μηχανικής Μάθησης σε μη επεξεργασμένα δεδομένα. Έτσι, αφαιρούμε το βάρος της προεπεξεργασίας από τους αλγορίθμους των σημάτων και το μεταφέρουμε στο μοντέλο, το οποίο πλέον μπορεί να κάνει χρήση όλης της διαθέσιμης πληροφορίας. Επιτρέπουμε στο μοντέλο να εξάγει από μόνο του τα απαραίτητα χαρακτηριστικά για την επεξεργασία του σήματος και δείχνουμε πως χωρίς καθόλου γνώση του τομέα (domain knowledge) μπορούμε να επιτύχουμε εξίσου καλά αποτελέσματα σε σχέση με τις παραδοσιακές τεχνικές.

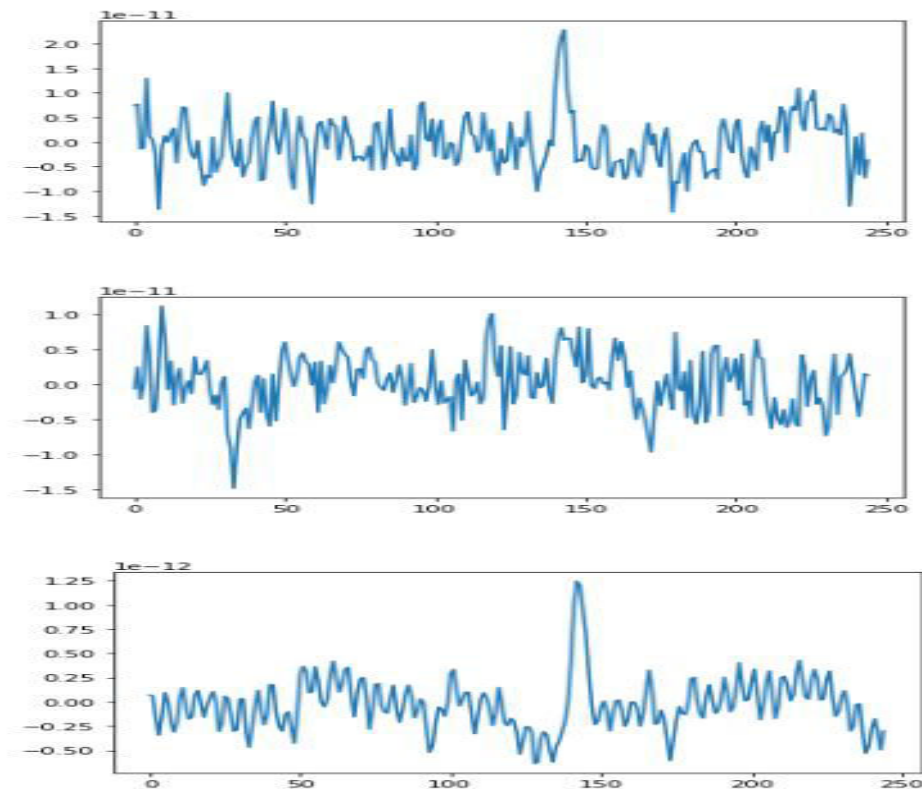
4 ΤΟ ΣΕΤ ΔΕΔΟΜΕΝΩΝ

Τα δεδομένα που είχαμε στη διάθεσή μας για εκπαίδευση αποτελούνταν από συνολικά 9414 δοκιμές (trials) που προέρχονταν από 16 διαφορετικά υποκείμενα, δηλαδή περίπου 580-590 δοκιμές ανά υποκείμενο. Η κάθε δοκιμή διαρκεί 1.5 δευτερόλεπτο και ξεκινάει 0.5 δευτερόλεπτο πριν από το ερέθισμα. Η εγκεφαλική δραστηριότητα μετريέται μέσω 306 αισθητήρων (που αποτελούν και τα διαφορετικά κανάλια) οι οποίοι μετρούν το μαγνητικό πεδίο που παράγεται λόγω των ρευμάτων των νευρώνων, με συχνότητα 1kHz. Οι αισθητήρες αυτοί βρίσκονται τοποθετημένοι σε τριάδες σε 102 σημεία του κεφαλιού του

υποκειμένου. Στη συνέχεια το σήμα υποδειγματοληπτείται με συχνότητα 250Hz και περνάει από ένα υπερηχητικό φίλτρο με συχνότητα αποκοπής το 1Hz. Αυτό έχει σαν αποτέλεσμα να παίρνουμε 375 δείγματα (samples) από κάθε ακολουθία (του κάθε αισθητήρα) με το συνολικό αποτέλεσμα για κάθε υποκείμενο να σχηματίζει έναν τρισδιάστατο πίνακα διαστάσεων 580(δοκιμές)X306(αισθητήρες)X375(δείγματα). Το αποτέλεσμα κάθε δοκιμής είναι δυαδικό με την τιμή 1 για την ύπαρξη προσώπου και την τιμή 0 για την ύπαρξη αλλοιωμένης εικόνας. Το τεστ δοκιμής (test set) αποτελούνταν από 4058 δοκιμές για 7 υποκείμενα, διαφορετικά από αυτά τα οποία εμφανίστηκαν στο training set για λόγους γενίκευσης.

4.1 Προεπεξεργασία

Πέρα από την επεξεργασία των σημάτων πριν έρθουν σε εμάς ήταν απαραίτητο να κάνουμε και εμείς κάποια προεργασία για να μπορέσουμε να επεξεργαστούμε τα δεδομένα περαιτέρω. Αρχικά, τα δεδομένα ήταν συμπιεσμένα και αποθηκευμένα σε μορφή *Matlab* v5.0 οπότε ήταν απαραίτητο να τα αποσυμπιέσουμε και να τα φέρουμε σε μορφή, τέτοια ώστε να μπορούν να αναγνωστούν μέσω της γλώσσας *Python* στην οποία έγινε και η υλοποίησή μας. Επίσης, αποφασίσαμε να αφαιρέσουμε το πρώτο μισό δευτερόλεπτο κάθε σήματος μιας και σε αυτή τη διάρκεια δεν υπάρχει κανένα ερέθισμα και ουσιαστικά δε φέρει πληροφορία. Έτσι, απομακρύνουμε περιττή πληροφορία ενώ ταυτόχρονα μειώσαμε και το φυσικό όγκο των δεδομένων (γεγονός το οποίο ήταν κρίσιμο λόγω της περιορισμένης μνήμης RAM). Ακόμα, τα σήματα αυτά ήταν της τάξης του 10^{-11} λόγω της ασθένειας των ηλεκτρικών σημάτων στον εγκέφαλο και έτσι ήταν απαραίτητο να τα φέρουμε σε μορφή στην οποία θα μπορούσαμε να τα επεξεργαστούμε χωρίς να έχουμε προβλήματα πεπερασμένης ακρίβειας και ταυτόχρονα θα βοηθήσουμε τα μοντέλα μας να συγκλίνουν. Έτσι, κανονικοποιήσαμε τα σήματα (Standard Scaling) με μέση τιμή μηδέν και τυπική απόκλιση 1. Τέλος, χωρίσαμε το σετ εκπαίδευσης σε training set, validation set και test set εκ των οποίων το πρώτο σετ δεδομένων χρησιμοποιήθηκε για εκπαίδευση, το δεύτερο σετ για την επιλογή των παραμέτρων (fine tuning) και τον έλεγχο γενίκευσης του μοντέλου και το τρίτο σετ για την αξιολόγηση των αποτελεσμάτων μας. Η διαμέριση έγινε με τυχαίο τρόπο (ανακατεύοντας τα δεδομένα για να μην επηρεάσει το μοντέλο η σειρά τους) ενώ δοκιμάστηκαν και διάφορα ποσοστά για το κάθε σετ. Παρακάτω, φαίνονται κάποια παραδείγματα από τα σήματα πριν την κανονικοποίηση.



Εικόνα 2. Παραδείγματα σημάτων από διαφορετικά κανάλια για μια δοκιμή

5 ΤΑ ΜΟΝΤΕΛΑ

Χρησιμοποιήσαμε διάφορα μοντέλα για την εύρεση της βέλτιστης λύσης και σε αυτό το σημείο θα τα παραθέσουμε, θα αναλύσουμε τον τρόπο λειτουργίας τους και θα εξετάσουμε τη λογική πίσω από τη χρήση τους.

5.1 Feed-Forward Neural Network

Το πρώτο μοντέλο που υλοποιήσαμε ήταν ένα Feed-Forward Νευρωνικό δίκτυο που αποτελούνταν από 3 πλήρως συνδεδεμένα επίπεδα (fully connected layers). Από την περιγραφή των δεδομένων νωρίτερα, είχαμε επισημάνει ότι τα δεδομένα για κάθε δοκιμή αποτελούν έναν δισδιάστατο πίνακα, οπότε για να τα επεξεργαστούμε με ένα feed-forward δίκτυο έπρεπε να τα μετατρέψουμε σε μονοδιάστατα, δηλαδή να απλώσουμε τη δεύτερη διάσταση στην πρώτη. Αυτή η μέθοδος είναι και η πιο αφελής, καθώς αγνοεί τελείως τη μορφοποίηση των δεδομένων και αντιμετωπίζει όλα τα διαφορετικά κανάλια ως ίδια.

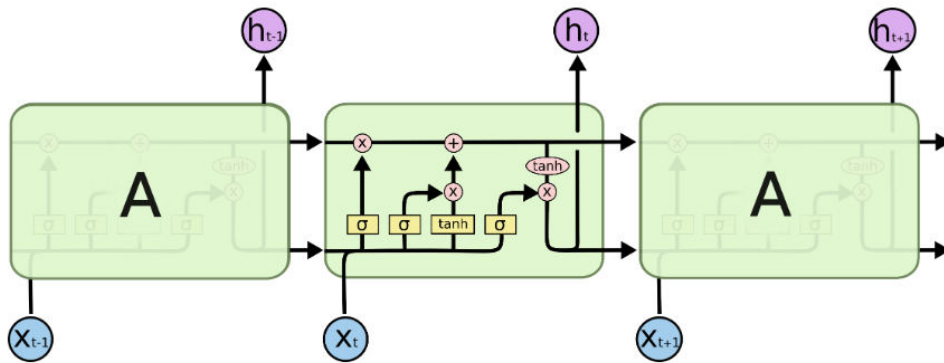
5.2 Long Short-Term Memory Network (LSTM)

Το LSTM αποτελεί ένα ειδικό αναδρομικό νευρωνικό δίκτυο το οποίο είναι κατάλληλο για τη μοντελοποίηση χρονοσειρών και γενικότερα σημάτων τα οποία εξελίσσονται στο χρόνο. Ανήκει στην κατηγορία των αναδρομικών νευρωνικών δικτύων, δηλαδή δικτύων τα οποία επαναλαμβάνονται για κάθε σημείο της ακολουθίας, μεταφέροντας όμως την πληροφορία μέσω της εσωτερικής δομής τους. Έτσι, μπορούν να μοντελοποιήσουν καλύτερα συσχετίσεις που υπάρχουν ακόμα και μακριά στο χρόνο. Αυτό το πετυχαίνουν διατηρώντας κομμάτια της πληροφορίας από τα προηγούμενα σημεία, ξεχνώντας πληροφορία η οποία δεν είναι χρήσιμη και τελικά ανανεώνοντας σε κάθε βήμα την πληροφορία που διατηρείται στο χρόνο, ελαχιστοποιώντας έτσι το σφάλμα. Πιο συγκεκριμένα όπως φαίνεται και από το παρακάτω σχήμα ουσιαστικά κάθε μονάδα του δικτύου αυτού αποτελείται από τέσσερα επίπεδα:

- τη κατάσταση (cell state) που διατρέχει όλες τις μονάδες και ουσιαστικά αποτελεί την κοινή πληροφορία που ρέει μεταξύ των μονάδων,
- την πύλη (forget gate), που ανάλογα με τα δεδομένα εισόδου και την προηγούμενη κρυφή κατάσταση (hidden state) καθορίζει (με τη χρήση σιγμουειδούς συνάρτησης ενεργοποίησης) πόση από την προηγούμενη πληροφορία πρέπει να «ξεχάσει»,
- την πύλη εισόδου (input gate), η οποία ανάλογα με τα δεδομένα εισόδου και την προηγούμενη κρυφή κατάσταση καθορίζει πόση από τη νέα πληροφορία (είσοδος) χρειάζεται να «απομνημονεύσει» το δίκτυο και έπειτα φροντίζει να ανανεώσει ανάλογα την κατάσταση
- και τέλος την πύλη εξόδου (output gate) η οποία ανάλογα με τα δεδομένα εισόδου και την προηγούμενη κρυφή κατάσταση καθορίζει πόσο θα πρέπει να επηρεάσει η cell state τη νέα hidden state η οποία αποτελεί ουσιαστικά και την έξοδο του LSTM τη στιγμή i .

Στην υλοποίησή μας ενδιαφερόμαστε για την συνολική πληροφορία που υπάρχει στο σήμα, οπότε χρησιμοποιούμε την τελική hidden state του δικτύου h_T όπου T το μήκος της ακολουθίας. Θεωρούμε ότι αυτό το διάνυσμα περιέχει όλη την πληροφορία από την κωδικοποίηση της ακολουθίας. Επίσης, στην υλοποίηση μας χρησιμοποιούμε μια τροποποίηση του LSTM, το αμφίδρομο (bidirectional) LSTM. Αποτελείται ουσιαστικά από δύο διαφορετικά LSTMs: ένα το οποίο διατρέχει την ακολουθία από την αρχή προς το τέλος και ένα άλλο το οποίο διατρέχει την ακολουθία από το τέλος προς την αρχή. Αυτό έχει το πλεονέκτημα ότι έχουμε περισσότερη πληροφορία για την ακολουθία και μπορούμε να

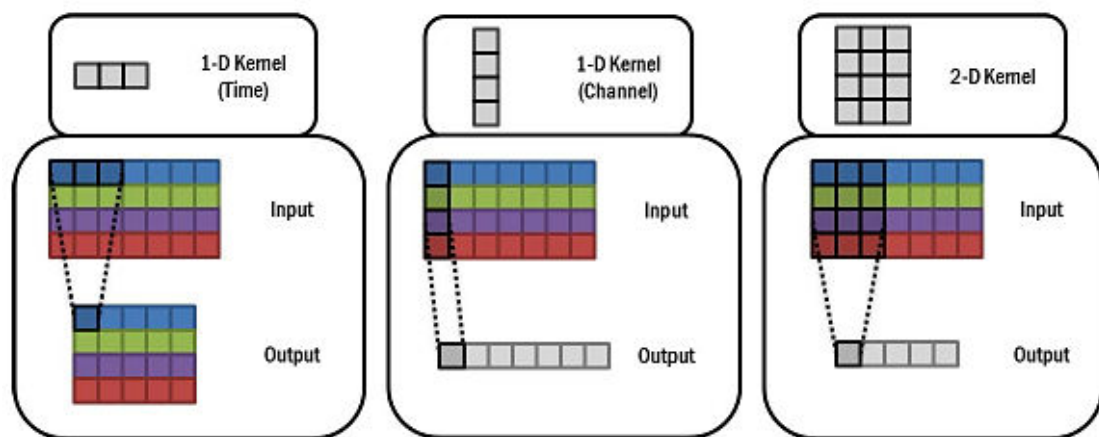
ανιχνεύσουμε καλύτερα συμβάντα τα οποία συμβαίνουν σε αυτή, όπως στην περίπτωση μας. Τέλος, εφαρμόσαμε και υλοποίηση με δύο επίπεδα LSTM (stacked LSTMs) εκμεταλλευόμενοι το γεγονός ότι όσο «βαθαίνει» ένα Νευρωνικό Δίκτυο τόσο περισσότερο είναι ικανό να μάθει πιο σύνθετες δομές στα δεδομένα.



Εικόνα 3. Εσωτερική δομή του LSTM

5.3 Convolutional Networks

Στην υλοποίησή μας, χρησιμοποιήσαμε ακόμα ένα είδος νευρωνικών δικτύων, τα Συνελικτικά Δίκτυα (Convolutional Networks). Τα δίκτυα αυτά, όπως διδαχτήκαμε και στο μάθημα υλοποιούν ουσιαστικά την πράξη της συνέλιξης και είχαν αρχικά δημιουργηθεί για εικόνες στις οποίες η πράξη της συνέλιξης ήταν ήδη διαδεδομένη για την επεξεργασία τους. Τα δίκτυα αυτά αρχικοποιούν τα βάρη του πυρήνα της συνέλιξης τυχαία και μέσω του αλγορίθμου της οπισθοδιάδοσης (backpropagation) «μαθαίνουν» τον βέλτιστο πυρήνα έτσι ώστε να ελαχιστοποιήσουν το σφάλμα. Στην περίπτωση μας χρησιμοποιούμε δυο παραλλαγές των Convolutional Networks: τα δίκτυα στα οποία ο πυρήνας είναι μονοδιάστατος, δηλαδή υλοποιείται η συνέλιξη στη μία διάσταση και τα δίκτυα εκείνα στα οποία ο πυρήνας είναι δισδιάστατος. Για την πρώτη περίπτωση δοκιμάσαμε δύο διαφορετικές υλοποιήσεις της μονοδιάστατης συνέλιξης: συνέλιξη στον άξονα του χρόνου και συνέλιξη στον άξονα των καναλιών. Συνολικά, οι διαφορετικές μας υλοποιήσεις με χρήση συνελικτικών δικτύων απεικονίζονται στο παρακάτω σχήμα.



Εικόνα 4. Διαφορετικές υλοποιήσεις της συνέλιξης στα μοντέλα μας

Η λογική πίσω από τη χρήση της μονοδιάστατης συνέλιξης είναι αυτή της εξαγωγής χαρακτηριστικών (feature extraction) από τα ανεπεξέργαστα δεδομένα, προσομοιώνοντας τις μεθόδους της επεξεργασίας σήματος για την εξαγωγή χρήσιμης πληροφορίας από τα δεδομένα. Αυτή η διαδικασία προηγείται της τροφοδότησής τους στο δίκτυο, διευκολύνοντας έτσι την ανάλυσή τους από αυτό. Για αυτό τον σκοπό, αρχικά χρησιμοποιήσαμε μονοδιάστατη συνέλιξη στο χρόνο έτσι ώστε το νέο, επεξεργασμένο πλέον, σήμα να φέρει πληροφορία από διάφορες τιμές που βρίσκονται κοντά στο χρόνο. Αυτό είναι θετικό και εφαρμόζεται συχνά κατά την ανάλυση χρονοσειρών καθώς οι πυρήνες αυτοί είναι ικανοί να κινηθούν και στο πεδίο της συχνότητας, καθώς συνδυάζουν τιμές από διάφορα χρονικά σημεία και έτσι προσφέρουν νέα χαρακτηριστικά στο δίκτυο. Επίσης, χρησιμοποιήσαμε και μονοδιάστατη συνέλιξη στο πεδίο των διάφορων καναλιών (αισθητήρων) έτσι ώστε να εκμεταλλευτούμε τη χωρική σχέση που υπάρχει μεταξύ των διαφόρων νευρώνων στον εγκέφαλο και κατά συνέπεια τη σχέση που υπάρχει μεταξύ των σημάτων των διαφόρων αισθητήρων. Αυτή η ιδέα έρχεται ως εναλλακτική της υλοποίησης αλγορίθμων που χρησιμοποιούν τη διαφορετική πληροφορία μεταξύ των καναλιών για την εξαγωγή καλύτερων χαρακτηριστικών (πχ. CSP). Μετά από τη μονοδιάστατη συνέλιξη τα νέα πλέον διανύσματα χαρακτηριστικών τροφοδοτούνται σε LSTMs επεκτείνοντας τη λογική που αναφέραμε στην προηγούμενη παράγραφο. Τέλος, η υλοποίηση με δισδιάστατη συνέλιξη αποτελεί την επέκταση των δυο προηγούμενων μεθόδων, εξάγοντας χαρακτηριστικά από τη διάσταση του χρόνου και τη διάσταση των καναλιών ταυτόχρονα. Αυτό μπορεί να συμβεί επειδή τα δεδομένα μας έχουν δύο διαστάσεις και πρακτικά μπορούμε να χρησιμοποιήσουμε μεθόδους από την επεξεργασία εικόνων με νευρωνικά δίκτυα.

5.3 Άλλες τεχνικές

Κατά την υλοποίηση των παραπάνω δικτύων χρησιμοποιήθηκαν, πέρα από τα μοντέλα και άλλες τεχνικές των νευρωνικών δικτύων που ως στόχο είχαν είτε τη βελτίωση του μοντέλου ή τη βελτίωση της δυνατότητας γενίκευσής του. Οι σημαντικότερες απ' αυτές περιγράφονται παρακάτω:

- Weight decay (αποσύνθεση βαρών): Η τεχνική αυτή προσθέτει έναν παραπάνω όρο στη συνάρτηση κόστους που περιορίζει τη μεγάλη αύξηση των βαρών και χρησιμοποιείται για την αποφυγή της υπερεκπαίδευσης (overfitting) στο σετ εκπαίδευσης. Είναι ισοδύναμη με τη χρήση μιας βοηθητικής L_2 συνάρτησης κόστους (L_2 regularization) με τη διαφορά ενός παράγοντα $\frac{1}{2}$.
- Dropout : Η μέθοδος αυτή χρησιμοποιείται κατά κόρον στα νευρωνικά δίκτυα και λειτουργεί επίσης σαν regulariser. Εν συντομία, κατά την εκπαίδευση των δεδομένων αφαιρούνται τυχαία (με βάση μια κατανομή Bernoulli) κάποιες συνάψεις των νευρώνων σε κάθε επανάληψη και ενημερώνονται αντίστοιχα κατά το backpropagation. Αυτό έχει σαν αποτέλεσμα το νευρωνικό δίκτυο να μην επηρεάζεται περισσότερο από κάποιους νευρώνες σε σχέση με τους υπόλοιπους και να αντλεί πληροφορία από όλα τα χαρακτηριστικά. Προφανώς, κατά το testing αυτή η μέθοδος απενεργοποιείται.
- Pooling : Αποτελεί μια ομάδα μεθόδων που στόχο έχουν την υποδειγματοληψία (subsampling) του διανύσματος εισόδου. Υπάρχουν διάφορες παραλλαγές (mean pooling, max pooling κλπ.) οι οποίες έχουν παρόμοιο τρόπο λειτουργίας: την επιλογή ενός δείγματος (η ενός συνδυασμού τους στην περίπτωση mean pooling) από μία γειτονία δειγμάτων. Αυτές οι τεχνικές χρησιμοποιούνται συνήθως μετά από συνελκτικά επίπεδα για μείωση των διαστάσεων.
- Batch Normalisation : Ακόμα μια μέθοδος regularization η οποία επεκτείνει την ιδέα της κανονικοποίησης, αυτή τη φορά σε επίπεδο νευρώνων των κρυφών επιπέδων. Η βασική λογική είναι η κανονικοποίηση των εισόδων κάθε νευρώνα με βάση τις τιμές τους μέσα σε ένα δείγμα δεδομένων (batch) για τη μείωση της αλλαγής της κατανομής εισόδου εξαιτίας της διαφορετικότητας των δεδομένων (internal covariate shift). Αποτρέπει την

υπερεκπαίδευση ενώ παράλληλα συμβάλει και στην ταχύτερη σύγκλιση του μοντέλου.

Τέλος, κατά την εκπαίδευση όλων των μοντέλων χρησιμοποιήθηκε η τεχνική του early stopping για την αποφυγή της υπερεκπαίδευσης μέσω του ελέγχου του σφάλματος σε κάθε εποχή στο validation set. Ως παράμετροι για αυτή τη μέθοδο χρησιμοποιήθηκαν patience=20 (δηλ. μέγιστος αριθμός εποχών που μπορεί να μη παρατηρηθεί μείωση στο σφάλμα του validation set) και max_epochs=50 (μέγιστος αριθμός εποχών που μπορεί να τρέξει το μοντέλο).

6 ΑΠΟΤΕΛΕΣΜΑΤΑ

Αφού αναλύσαμε στην προηγούμενη παράγραφο τον τρόπο λειτουργίας και τη λογική πίσω από τη χρησιμοποίηση των μοντέλων μας, σε αυτή την παράγραφο θα παρουσιάσουμε τα αποτελέσματά τους. Αρχικά θα αναφέρουμε τις παραμέτρους που χρησιμοποιήθηκαν για όλα τα μοντέλα. Οι παράμετροι αυτές βρέθηκαν από εμάς μετά από δοκιμές στο validation set.

- Για το Fully Connected δίκτυο χρησιμοποιήσαμε 3 layers με πλήθος νευρώνων $74970 (= 254 * 306)$, 200 και 20 αντίστοιχα, ενώ χρησιμοποιήσαμε και weight decay της τάξης του 0.1.
- Για το LSTM χρησιμοποιήσαμε αμφίδρομο LSTM, με 2 layers (stacked LSTMs), dropout 0.2 και hidden dimension 20. Επίσης, τροφοδοτούμε την έξοδο του bidirectional LSTM σε ένα fully connected layer με 40 νευρώνες.
- Για την υλοποίηση με τη μονοδιάστατη συνέλιξη στο χρόνο χρησιμοποιήσαμε 2 πυρήνες μεγέθους 3, με padding 1 και το LSTM της προηγούμενης υλοποίησης με hidden state dimension 40 και dropout 0.3
- Για την υλοποίηση με τη μονοδιάστατη συνέλιξη στα κανάλια χρησιμοποιήσαμε 1 πυρήνα μεγέθους 5, με padding 3 και το LSTM της προηγούμενης υλοποίησης.
- Για το δισδιάστατο συνελκτικό δίκτυο χρησιμοποιήσαμε 4 επίπεδα, καθένα από τα οποία αποτελούνταν από ένα συνελκτικό επίπεδο, ένα επίπεδο batch normalization, ένα επίπεδο max pooling και συνάρτηση ενεργοποίησης ReLU. Ο αριθμός πυρήνων για κάθε τέτοιο επίπεδο ήταν αντίστοιχα 16, 8, 4 και 2. Τέλος, μετά από το τελευταίο συνελκτικό επίπεδο χρησιμοποιήσαμε και ένα fully connected layer.

Στον παρακάτω πίνακα συνοψίζονται τα αποτελέσματα που λάβαμε για τις παραπάνω παραμέτρους. Τα αποτελέσματα προέκυψαν από 5 διαφορετικές δοκιμές των μοντέλων μας (συμπεριλαμβανομένου και του χωρισμού του dataset) για την εξαγωγή ασφαλέστερων συμπερασμάτων για την ακρίβεια τους. Όλα τα μοντέλα εκπαιδεύτηκαν με batch size 64 και με συνάρτηση κόστους την Binary Cross Entropy. Παρουσιάζονται τα αποτελέσματα στο validation και το test set.

model	Accuracy(test val)	
DNN	64.6±4.56%	64.03±5.97%
LSTM	56.4±0.49%	57.49±0.97%
LSTM+1dConv (in time)	70.15±0.56%	71.39±0.82%
LSTM+1dConv (in space)	67.44±0.63%	68.16±0.76%
CNN	68.65±0.65%	69.32±0.78%

7 ΣΥΜΠΕΡΑΣΜΑΤΑ

Αρχικά παρατηρούμε ότι το DNN έχει πολύ μεγάλη διακύμανση, γεγονός που οφείλεται στις πολλές παραμέτρους του δικτύου. Επειδή συμπύξαμε τις δύο διαστάσεις των δεδομένων σε μία, το διάνυσμα εισόδου ήταν τεράστιο και το δίκτυο οδηγούνταν γρήγορα στο overfit εξαιτίας αυτού. Για αυτό το λόγο χρησιμοποιήσαμε και weight decay μεγαλύτερο από τα άλλα μοντέλα, αλλά και πάλι δε μπορέσαμε να μειώσουμε το variance του μοντέλου. Η χρήση ενός τέτοιου δικτύου για την επεξεργασία δισδιάστατων δεδομένων ήταν εξ αρχής άτοπη και χρησιμοποιήθηκε αποκλειστικά για την ύπαρξη ενός βασικού, απλού μοντέλου ως μέτρου σύγκρισης (baseline).

Επίσης, παρατηρούμε ότι το LSTM χωρίς κάποιο άλλο υποδίκτυο, έχει πολύ μικρή ακρίβεια, γεγονός που οφείλεται στην αδυναμία του δικτύου να κάνει από μόνο του το απαραίτητο feature extraction για να επεξεργαστεί βέλτιστα τα δεδομένα. Αυτό ήταν αναμενόμενο και είναι και ο λόγος που εισάγουμε τα παρακάτω μοντέλα.

Ακόμα, παρατηρούμε ότι την καλύτερη απόδοση έχει το LSTM με 1D Convolution στη διάσταση του χρόνου. Ο πιθανότερος λόγος για αυτή την απόδοση είναι ότι οι πυρήνες λειτουργούν πολύ καλά ως feature extractors στο πεδίο του χρόνου, επειδή μπορούν να μεταβούν και στο πεδίο της συχνότητας, που είναι σύνηθες στην επεξεργασία σημάτων.

Ακόμα, το LSTM με 1D Convolution στη διάσταση των καναλιών έχει καλή απόδοση αλλά όχι τόσο καλή όσο θα περιμέναμε, πιθανώς εξαιτίας της απουσίας της σωστής διάταξης στα κανάλια που δεν ευνοεί την εξαγωγή χρήσιμης πληροφορίας από γειτονικά κανάλια.

Τέλος, το CNN λειτουργεί και αυτό πολύ καλά ώντας ένας συνδυασμός των συνελίξεων στο χρόνο και στο χώρο των καναλιών.

8 ΕΠΕΚΤΑΣΕΙΣ

Μια απλή και ταυτόχρονα χρήσιμη πληροφορία που έλειπε από τα μοντέλα μας είναι η σχετική θέση των αισθητήρων στον εγκέφαλο. Αυτή η πληροφορία θα μας έδινε τη δυνατότητα να αναδιατάξουμε τα κανάλια με τρόπο τέτοιο, ώστε κανάλια τα οποία σχετίζονται χωρικά να είναι κοντά, έτσι ώστε να παίρνουμε τη μέγιστη πληροφορία από τους πυρήνες της συνέλιξης. Αυτή η προσθήκη θα βελτίωνε τα μοντέλα που κάνουν χρήση αυτής της χωρικής πληροφορίας, δηλαδή αυτά που χρησιμοποιούν συνελκτικά επίπεδα σε μία και σε δύο διαστάσεις.

Ακόμα, για το δισδιάστατο συνελκτικό δίκτυο θα μπορούσαμε να χρησιμοποιήσουμε ένα προεκπαιδευμένο σε εικόνες δίκτυο και να κάνουμε χρήση του transfer learning για να προσαρμόσουμε αυτό το μοντέλο στο πρόβλημά μας. Αυτή είναι μια τακτική που χρησιμοποιείται ευρέως στα προβλήματα επεξεργασίας εικόνων (object detection, segmentation κλπ.) καθώς υπάρχουν προεκπαιδευμένα μοντέλα τα οποία έχουν πολύ υψηλή ακρίβεια (VGG, Inception, ResNet κλπ.) και μπορούν να χρησιμοποιηθούν ως feature extractors για το υπόλοιπο δίκτυο. Έτσι, με χρήση αυτών των αρχιτεκτονικών και με κατάλληλη εκπαίδευση (fine tuning) στο πρόβλημά μας θα μπορούσαμε να εκμεταλλευτούμε την τεράστια ακρίβειά τους για να βελτιώσουμε τα δικά μας αποτελέσματα.

Τέλος, σε αυτή την εργασία πειραματιστήκαμε με διάφορες μεθοδολογίες των Νευρωνικών Δικτύων για την επεξεργασία εγκεφαλικών σημάτων χωρίς πρότερη γνώση του τομέα για να εξετάσουμε κατά πόσο αυτό θα ήταν εφικτό. Όπως φάνηκε από τα αποτελέσματα αυτή η μέθοδος λειτούργησε αρκετά καλά, καταδεικνύοντας τη δύναμη των Νευρωνικών Δικτύων που δικαιολογεί και τη χρήση τους σε όλο και περισσότερα προβλήματα. Παρ' όλα αυτά μια μελλοντική επέκταση θα ήταν η ενσωμάτωση της γνώσης του τομέα της Επεξεργασίας Σημάτων σε αυτά τα μοντέλα και την εξέταση της διαφοράς στην απόδοση που θα προσέφερε αυτή η επιπλέον γνώση.

9 ΑΝΑΦΟΡΕΣ

- Benjamin Blankertz, R. T.-R. (2008). Optimizing Spatial Filters for Robust EEG Single-Trial Analysis. *IEEE SIGNAL PROCESSING MAGAZINE*.
- Kai Keng Ang, C. G. (2008). Filter Bank Common Spatial Pattern (FBCSP) in brain-computer interface.
- Kai Keng Ang, Z. C. (2012). Filter bank common spatial pattern algorithm on BCI competition IV Datasets 2a and 2b. *Frontiers in Neuroscience*.
- Siavash Sakhavi, C. G. (2018). Learning Temporal Information for Brain-Computer Interface Using Convolutional Neural Networks. *TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS*,.