

Ανάλυση ευαισθησίας σε αναδρομικό νευρωνικό δίκτυο εκπαιδευμένο για αναγνώριση συναισθήματος

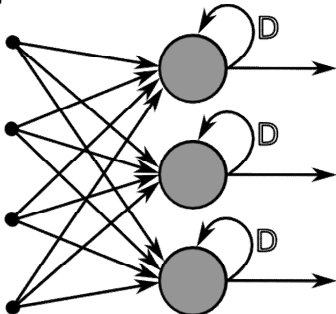
Σαρλίν Χονδρού, Μηνάς Περτσελάκης, Γιώργος Καρυδάκης, Κώστας Καρπούζης, Στέφανος Κόλλιας
Εργαστήριο Επεξεργασίας Εικόνων, Βίντεο και Πολυμέσων, ΗΜΜΥ, ΕΜΠ
[gcari, kkarrou, stefanos]@image.ece.ntua.gr

Abstract

Στην παρούσα εργασία εξετάζεται η αναγνώριση της συναισθηματικής κατάστασης ενός ανθρώπου από τα φυσικά χαρακτηριστικά του προσώπου του και την ομιλία του. Η διαδικασία της αναγνώρισης πραγματοποιείται με την βοήθεια ενός αναδρομικού νευρωνικού δικτύου που βασίζεται στην αρχιτεκτονική Elman. Επιπλέον, μελετάται η λειτουργία και η εκπαίδευση τέτοιου τύπου νευρωνικών δικτύων, όπως επίσης ο τρόπος με τον οποίο μπορούν να αξιολογηθούν και στη συνέχεια να αφαιρεθούν τα χαρακτηριστικά εισόδου τους με χρήση στατιστικής μεθόδου, γνωστής ως ανάλυση ευαισθησίας. Τα παραγόμενα αποτελέσματα από την χρήση αυτής της μεθόδου, οδηγούν σε σημαντικές πληροφορίες για την επίδραση κάθε εισόδου στο σύστημα ενώ παράλληλα, επιτυγχάνουν να μειώσουν την διάσταση του προβλήματος, άρα και την πολυπλοκότητά του, καθώς και να βελτιώσουν την ταχύτητα και την απόδοση του δικτύου.

1. Εισαγωγή

Τα τεχνητά αναδρομικά (ή ανατροφοδοτούμενα) νευρωνικά δίκτυα [3] αποτελούν την εξέλιξη των τεχνητών νευρωνικών δικτύων [2] τα οποία, με την σειρά τους, επιχειρούν να προσομοιώσουν την λειτουργία των φυσικών νευρωνικών δικτύων.



Σχήμα 1: Αναδρομικό Νευρωνικό Δίκτυο

Χρησιμοποιείται ένα δίκτυο τριών επιπέδων στο οποίο έχει προστεθεί ένα σύνολο από μονάδες περιεχομένου (context units) στο επίπεδο εισόδου. Υπάρχουν συνδέσεις από το μεσαίο (κρυφό) επίπεδο σε αυτές τις μονάδες περιεχομένου οι οποίες έχουν σταθερό βάρος ίσο με 1. Σε κάθε βήμα, η είσοδος προωθείται προς το επόμενο επίπεδο (feed-forward) κι έπειτα εφαρμόζεται ένας κανόνας εκμάθησης. Οι σταθερές συνδέσεις προς τα πίσω καταλήγουν στις μονάδες περιεχομένου κρατώντας ένα αντίγραφο των προηγούμενων τιμών από τις κρυφές μονάδες. Έτσι, το δίκτυο μπορεί να διατηρήσει μία κατάσταση, έτσι ώστε να έχει την ικανότητα να φέρνει εις πέρας εργασίες όπως πρόβλεψη ακολουθίας που είναι πέρα από τις δυνατότητες ενός απλού πολυεπίπεδου νευρωνικού δικτύου.

Η ανάλυση ευαισθησίας υπέδειξε τις εισόδους με τη μικρότερη επίδραση στην εξαγωγή αποτελέσματος κατά την εκπαίδευση του δικτύου. Η απενεργοποίηση αυτών των εισόδων κατά την δοκιμή μας οδήγησε στη μείωση της διάστασης του προβλήματος. Το αποτέλεσμα αυτό μας φέρνει ένα βήμα πιο κοντά στην αντιμετώπιση ενός από τα προβλήματα των νευρωνικών δικτύων: η αύξηση της διάστασης του προβλήματος ακολουθείται από την εκθετική αύξηση της ανάγκης για κρυφούς κόμβους, άρα και της πολυπλοκότητας του συστήματος. Οπότε μειώνοντας τις εισόδους, μειώσαμε την πολυπλοκότητα με επιπρόσθετες συνέπειες τη μεγαλύτερη ταχύτητα και την καλύτερη απόδοση στην αναγνώριση της συναισθηματικής κατάστασης. Η ανάλυση ευαισθησίας υπέδειξε ακόμα τις εισόδους που επηρεάζουν δραστηκότερα την κάθε μια από τις εξόδους του νευρωνικού δικτύου.

Οι τεχνικές που προαναφέρθηκαν, αναδρομικά νευρωνικά δίκτυα και ανάλυση ευαισθησίας, συνδυάστηκαν και βρήκαν εφαρμογή στην αναγνώριση του συναισθήματος του ομιλητή μιας μαγνητοσκοπησης καθώς και στην εξαγωγή συμπερασμάτων για το μέγεθος της επιρροής των χαρακτηριστικών του προσώπου του και της φωνής του στην αξιολόγηση της συναισθηματικής του κατάστασης.

2. Λειτουργία αναδρομικού νευρωνικού δικτύου Elman

Θα εξετάσουμε την λειτουργία ανατροφοδοτούμενου νευρωνικού δικτύου ενός επιπέδου αλλά η λογική επαληθεύεται και για δίκτυα περισσότερων κρυφών επιπέδων. Θεωρούμε:

- $x_i(t)$ την i -οστή είσοδο του δικτύου
- $y_j(t)$ την έξοδο του j -κρυφού νευρώνα
- $y_h(t-1)$ την προηγούμενη χρονικά έξοδο του h -κρυφού νευρώνα
- $y_k(t)$ την k -οστή έξοδο του δικτύου

Κατά την εκπαίδευσή του, το δίκτυο κατασκευάζει τους πίνακες βαρών V , U και W τα στοιχεία των οποίων είναι αντίστοιχα:

- v_{ji} το βάρος ανάμεσα στην i -είσοδο και τον j -κρυφό νευρώνα
- u_{jh} το βάρος ανάμεσα στην προηγούμενη χρονικά έξοδο του h -κρυφού νευρώνα και την παρούσα είσοδο του j -κρυφού νευρώνα
- w_{kj} το βάρος ανάμεσα στην έξοδο του j -κρυφού νευρώνα και την k -οστή έξοδο του δικτύου.

Για λόγους προφύλαξης του δικτύου κάθε νευρώνας των δύο τελευταίων επιπέδων, λαμβάνει και μια σταθερή είσοδο που ονομάζεται bias. Στο δίκτυο που μελετάμε έχουμε:

- J_j το bias του j -κρυφού νευρώνα
- J_k το bias της k -εξόδου.

Για να δημιουργηθεί η συνολική είσοδος κάθε νευρώνα αθροίζονται τα γινόμενα της εκάστοτε εισόδου με το αντίστοιχο βάρος και προστίθεται και το bias. Το αποτέλεσμα αυτό μπαίνει ως όρισμα στην συνάρτηση ενεργοποίησης του επιπέδου. Οι συναρτήσεις ενεργοποίησης των δύο τελευταίων επιπέδων μπορεί να είναι ίδιες ή διαφορετικές. Ορίζουμε:

- f την συνάρτηση ενεργοποίησης του κρυφού επιπέδου
- g την συνάρτηση ενεργοποίησης της εξόδου.

Οι εξισώσεις που δίνουν την έξοδο του κρυφού επιπέδου και την έξοδο του δικτύου είναι αντίστοιχα:

$$y_j(t) = f\left(\sum_i v_{ji} x_i(t) + \sum_h u_{jh} y_h(t-1) + J_j\right)$$

$$y_k(t) = g\left(\sum_j w_{kj} y_j(t) + J_k\right)$$

Συγκεντρωτικά, όλα τα παραπάνω παρουσιάζονται στο Σχήμα 2.

Η δημιουργία αναδρομικού νευρωνικού δικτύου προϋποθέτει τον προσδιορισμό ορισμένων χαρακτηριστικών:

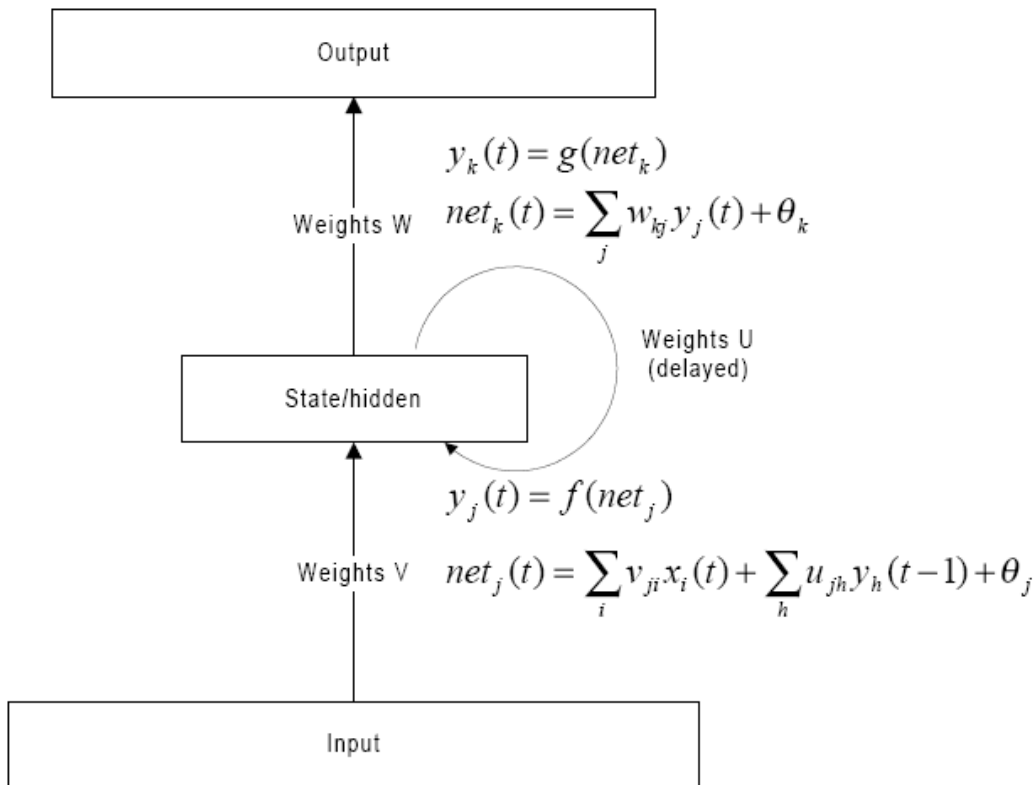
- Την αρχιτεκτονική (δομή και πλήθος κόμβων)
- Τις δύο συναρτήσεις ενεργοποίησης

Στόχος της εκπαίδευσης του νευρωνικού δικτύου είναι είναι η οριστικοποίηση της συνάρτησης μεταφοράς του μέσω του καθορισμού των πινάκων βαρών. Για να γίνει αυτό, θα πρέπει να ορίσουμε τον πίνακα εισόδων και τον πίνακα επιθυμητών εξόδων. Άλλες, απαραίτητες για την εκπαίδευση, παράμετροι προς προσδιορισμό είναι:

- Το πλήθος των εποχών (φάσεις εκπαίδευσης)
- Ο ρυθμός μάθησης
- Η ορμή

Οι ονομασίες των παραμέτρων είναι από μόνες τους επεξηγηματικές, με εξαίρεση την ορμή την οποία και θα εξηγήσουμε. Η ορμή προσθέτει ένα τμήμα της προηγούμενης ανανέωσης βάρους στην τωρινή. Αυτό βοηθάει το σύστημα να συνεχίσει την πορεία του προς την ελαχιστοποίηση του σφάλματος, στην περίπτωση που έχει "φυλακιστεί" σε κάποιο τοπικό ελάχιστο. Η μεθοδολογία που ακολουθείται συνήθως για την εκπαίδευση ενός νευρωνικού δικτύου είναι η εξής:

Ένα μέρος του συνόλου δεδομένων μας (data set) χρησιμοποιείται για την εκπαίδευση του δικτύου (train set) ενώ το υπόλοιπο χρησιμοποιείται για να ελεγχθεί η αποδοτικότητα του δικτύου σε άγνωστα δεδομένα (test set). Στην συνέχεια, εφόσον έχει αποφασιστεί η αρχιτεκτονική του δικτύου, γίνονται δοκιμές αλλάζοντας τις παραμέτρους εκπαίδευσης έως ότου να βρεθεί ο συνδυασμός που οδηγεί στην καλύτερη απόδοση. Στόχος είναι το νευρωνικό να εκπαιδευτεί καλά, χωρίς όμως να φτάσει στην υπερεκπαίδευση. Η τελευταία αφορά συνήθως την κλάση εξόδου που εμφανίζεται σε πολύ μεγάλο ποσοστό σε σχέση με τις άλλες στο σύνολο εκπαίδευσης. Τότε έχουμε το φαινόμενο κατά το οποίο το σύστημα αναγνωρίζει παντού την έξοδο αυτή. Αυτό, σε συνδυασμό με το γεγονός ότι η συγκεκριμένη έξοδος συναντάται πολύ συχνά στο σύνολο εκπαίδευσης, έχει ως αποτέλεσμα ένα αρκετά καλό ποσοστό απόδοσης το οποίο όμως είναι πλασματικό. Αυτή ήταν και μία από τις δυσκολίες που συναντήσαμε στην πειραματική μας μελέτη.



Σχήμα 2: Λειτουργία Αναδρομικού Νευρωνικού Δικτύου

3. Ανάλυση ευαισθησίας

Η εφαρμογή αλγορίθμων εξαγωγής κανόνων στα φυσικά προβλήματα είναι η βασική δυσκολία στην προσπάθεια σύνθεσης αντιπροσωπευτικού συνόλου δεδομένων. Οι ερευνητές συχνά δυσκολεύονται να εντοπίσουν εκείνες τις μεταβλητές που επηρεάζουν πραγματικά την έξοδο του προβλήματος.

Τα νευρωνικά δίκτυα μαθαίνουν από τα δεδομένα εισόδου (σε συνδυασμό με τα δεδομένα εξόδου). Όταν, όμως, ο όγκος δεδομένων είναι μεγάλος, είναι δύσκολο να δει κανείς τις σχέσεις και αλληλεπιδράσεις μέσα σε αυτόν. Συνεπώς, μπορεί έτσι να συμπεριληφθούν περιττά δεδομένα εισόδου ή να παραλειφθούν μία ή περισσότερες σημαντικές παράμετροι.

Η ανάλυση ευαισθησίας εφαρμοσμένη σε ένα εκπαιδευμένο νευρωνικό δίκτυο είναι ένα αποδοτικό εργαλείο [6] για τον εντοπισμό των μεταβλητών εισόδων [1] που επηρεάζουν σημαντικά οποιαδήποτε από τις δυνατές εξόδους καθώς και αυτών με την μικρότερη επιρροή στο τελικό αποτέλεσμα. Οι εισοδοί που αποδεικνύεται ότι δεν επηρεάζουν την έξοδο μπορούν να αφαιρεθούν ώστε να μειωθεί η διάσταση του

προβλήματος με αποτέλεσμα να μειωθεί η πολυπλοκότητα του προβλήματος μας.

Επειδή η πειραματική μελέτη που ακολουθεί σε επόμενο κεφάλαιο ασχολείται με το αναδρομικό δίκτυο Elman, στο κεφάλαιο αυτό θα επικεντρωθούμε στην ανάλυση ευαισθησίας εφαρμοσμένη στο νευρωνικό αυτό. Παρακάτω αναλύουμε την τεχνική που αναπτύχθηκε για μη αναδρομικό νευρωνικό δίκτυο [5], έχοντας κάνει τις απαραίτητες προσαρμογές για να μπορέσει να εφαρμοστεί σε ανατροφοδοτούμενο δίκτυο.

Για το νευρωνικό δίκτυο τριών επιπέδων έχουμε:

- $\mathbf{z} = (z_1, \dots, z_i, \dots, z_J)$,
- $\mathbf{y} = (y_1, \dots, y_j, \dots, y_J)$
- $\mathbf{o} = (o_1, \dots, o_k, \dots, o_K)$

που εκπροσωπούν το επίπεδο εισόδου, το κρυμμένο επίπεδο και το επίπεδο εξόδου αντίστοιχα. Θεωρούμε το ζεύγος εκπαίδευσης $p = (\mathbf{z}^{(p)}, \mathbf{t}^{(p)})$ όπου $i = (t_1, \dots, t_k, \dots, t_K)$ οι επιθυμητές εξόδου. Για δεδομένο ζεύγος p ορίζεται η ευαισθησία $s_{ki}^{(p)}$ μιας εκπαιδευμένης εξόδου o_k ως προς μία είσοδο z_i ως εξής:

$$S_{ki}^{(p)} = \frac{\partial o_k}{\partial z_i} = o_k \sum_{j=1}^J w_{kj} \frac{\partial y_j}{\partial z_i} = o'_k \sum_{j=1}^J w_{kj} y'_j v_{ji} \quad (1)$$

όπου:

y_j η έξοδος του j -κρυφού νευρώνα, o'_k η παράγωγος της συνάρτησης ενεργοποίησης του επιπέδου εξόδου:

$$o_k = f\left(\sum_{j=1}^J w_{kj} y_j + b_k\right) \quad (2)$$

και y'_j η παράγωγος της συνάρτησης ενεργοποίησης του κρυφού επιπέδου:

$$y_j = f\left(\sum_{i=1}^I v_{ji} z_i + \sum_{h=1}^{J,H} u_{hj} y_j(t-1) + b_j\right) \quad (3)$$

όπου: w_{kj} είναι το βάρος ανάμεσα στον κρυφό νευρώνα y_j και την έξοδο o_k , b_k το bias της k -εξόδου, v_{ji} το βάρος ανάμεσα στην είσοδο z_i και τον κρυφό νευρώνα y_j , u_{hj} η έξοδος του κρυφού νευρώνα j από την προηγούμενη χρονική στιγμή, u_{hj} το βάρος ανάμεσα στην έξοδο του κρυφού νευρώνα j την προηγούμενη χρονική στιγμή και την είσοδο του κρυφού νευρώνα h και b_j το bias του κρυφού νευρώνα j .

Η εξίσωση (1) παριστάνει την ευαισθησία της εξόδου o_k σε σχέση με την είσοδο z_i για το ζεύγος P . Για το ζεύγος αυτό, ο πίνακας ευαισθησίας $S^{(p)}$ που αποτελείται από στοιχεία $S_{ki}^{(p)}$ ορίζεται ως εξής:

$$S^{(p)} = O' W Y' V \quad (4)$$

όπου $W^{(K \times J)}$ και $V^{(J \times I)}$ είναι οι πίνακες βαρών επιπέδου εξόδου και κρυφού επιπέδου αντίστοιχα και $O'^{(K \times K)}$ και $Y'^{(J \times J)}$ ορίζονται ως:

$$\begin{aligned} O' &= \text{diag}(o'_1, \dots, o'_K) \\ Y' &= \text{diag}(y'_1, \dots, y'_J) \end{aligned} \quad (5)$$

Η εξίσωση (4) δίνει τον πίνακα ευαισθησίας για ένα συγκεκριμένο ζεύγος εκπαίδευσης P . Όμως, κάθε ζεύγος P παράγει έναν πίνακα ευαισθησίας $S^{(p)}$. Για να εφαρμοστεί η ανάλυση ευαισθησίας ο πίνακας ευαισθησίας $S^{(p)}$ πρέπει να υπολογιστεί για ολόκληρο το σύνολο δεδομένων εισόδου.

Ο μέσος τετραγωνικός πίνακας ευαισθησίας S_{avg} ορίζεται ως εξής:

$$S_{ki,avg} = \sqrt{\frac{\sum_{p=1}^P [S_{ki}^p]^2}{p}} \quad (6)$$

Για να γίνει σωστή σύγκριση ανάμεσα στις εισόδους, είναι απαραίτητο να κανονικοποιηθούν εισοδοί και εξοδοί στην ίδια κλίμακα. Αυτό γίνεται με τον τύπο:

$$S_{ki,avg} = S_{ki,avg} \frac{(\max_{p=1,\dots,P} \{z_i^{(p)}\} - \min_{p=1,\dots,P} \{z_i^{(p)}\})}{(\max_{p=1,\dots,P} \{o_k^{(p)}\} - \min_{p=1,\dots,P} \{o_k^{(p)}\})} \quad (7)$$

Χρησιμοποιώντας την σχέση (6) η διάσταση του διανύσματος εισόδου \underline{z} μπορεί να μειωθεί αφαιρώντας εκείνες τις εισόδους z_i που δεν επηρεάζουν πολύ τις εξόδους. Ορίζεται η σπουδαιότητα Φ_i της εισόδου z_i για όλες τις εξόδους ως:

$$\Phi_i = \max_{k=1,\dots,K} \{S_{ki,avg}\} \quad (8)$$

Επειτα, τα στοιχεία του διανύσματος $\underline{\Phi}$ κατανομούνται σε αύξουσα σειρά έτσι ώστε τα πρώτα στοιχεία του να αντιστοιχούν στις εισόδους με την μικρότερη επίδραση στις εξόδους.

3.1 Αλγόριθμος μείωσης δεδομένων εισόδου

Παρακάτω παρουσιάζονται δύο αλγόριθμοι που στοχεύουν στη μείωση χαρακτηριστικών εισόδων οι οποίοι έχουν κοινά τα τέσσερα πρώτα βήματα και διαφοροποιούνται στην συνέχεια. (Ο αλγόριθμος που θα χρησιμοποιηθεί στην δική μας πειραματική μελέτη είναι ο δεύτερος).

3.1.1 Αλγόριθμος 1

Βήμα 1: Εκπαίδευση νευρωνικού δικτύου με τα αρχικά δεδομένα εισόδου.

Βήμα 2: Υπολογισμός όλων των ευαισθησιών μεταξύ εισόδων και εξόδων για κάθε ζεύγος εκπαίδευσης χρησιμοποιώντας την (4).

Βήμα 3: Υπολογισμός του πίνακα ευαισθησίας για όλες τις παραμέτρους εκπαίδευσης με χρήση της (6).

Βήμα 4: Υπολογισμός της σπουδαιότητας της κάθε εισόδου για όλες τις εξόδους μέσω της (8).

Βήμα 5: Αφαίρεση των παραμέτρων εισόδου με την μικρότερη σπουδαιότητα.

Βήμα 6: Επανεκπαίδευση του νευρωνικού με τις εναπομένουσες εισόδους.

Βήμα 7: Επανάληψη των βημάτων 1 έως 6 έως ότου να μην υπάρχουν άλλες εισοδοί με μικρή επιρροή στην έξοδο

3.1.2 Αλγόριθμος 2

Βήμα 1: Εκπαίδευση νευρωνικού δικτύου με τα αρχικά δεδομένα εισόδου.

Βήμα 2: Υπολογισμός όλων των ευαισθησιών μεταξύ εισόδων και εξόδων για κάθε ζεύγος εκπαίδευσης χρησιμοποιώντας την (4).

Βήμα 3: Υπολογισμός του πίνακα ευαισθησίας για όλες τις παραμέτρους εκπαίδευσης με χρήση της (6).

Βήμα 4: Υπολογισμός της σπουδαιότητας της κάθε εισόδου για όλες τις εξόδους μέσω της (8).

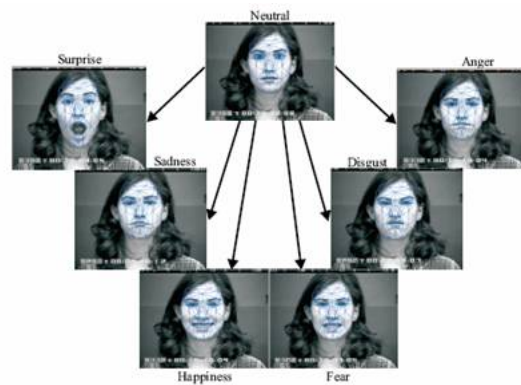
Βήμα 5: Έλεγχος απόδοσης του νευρωνικού χρησιμοποιώντας μόνο τα σπουδαιότερα χαρακτηριστικά.

Η πολύ σημαντική ιδιότητα των αναδρομικών δικτύων, δηλαδή η μνήμη, αντιπροσωπεύεται από τον δεύτερο όρο του δεύτερου μέλους της εξίσωσης (3). Με τον τρόπο αυτό, ενσωματώσαμε την ανατροφοδότηση σε μία μέθοδο της στατιστικής, όπως είναι η ανάλυση ευαισθησίας, της οποίας τις ιδιότητες θα αξιοποιήσουμε για την αναγνώριση συναισθηματικής κατάστασης.

4. Αναγνώριση ανθρώπινου συναισθήματος

Πόσο καλά, όμως, πρέπει να αναγνωρίζει ένας υπολογιστής το ανθρώπινο συναίσθημα για να θεωρηθεί ευφυής [7]; Πρέπει να σημειώσουμε ότι κανένας άνθρωπος δεν μπορεί να κατανοεί απόλυτα όλα τα συναισθήματα καθώς και ότι κάποιες φορές οι άνθρωποι δεν μπορούν να κατανοήσουν ούτε καν τα δικά τους συναισθήματα. Κανένας τρόπος αναγνώρισης συναισθήματος δεν είναι αλάθητος, όμως μερικές πτυχές των εσωτερικών συναισθηματικών διαβαθμίσεων παραμένουν κρυφές, ειδικά αν το άτομο το επιδιώκει ή αν επιλέγει να τα «μεταμφιέσει». Αυτό που μπορεί να δει κάποιος είναι αυτό που μπορεί να παρατηρηθεί και να εκλογικευθεί, και αυτό συνοδευόμενο πάντα με κάποια αμφιβολία. Παρόλα αυτά, οι άνθρωποι αναγνωρίζουν τα συναισθήματα ο ένας του άλλου, αρκετά καλά ώστε να μπορούν να προσαρμόζονται ανάλογα. Ο απώτερος στόχος, επομένως, είναι να αποκτήσουν οι υπολογιστές ικανότητες αναγνώρισης συναισθημάτων ανάλογες με αυτές των ανθρώπων.

Τα δεδομένα μας εξήχθησαν από μία μαγνητοσκοπημένη ομιλία και χωρίζονται σε οπτικά και ακουστικά. Τα οπτικά δίνουν πληροφορίες για τη μορφή των χαρακτηριστικών του προσώπου ανάλογα με την έκφραση (Σχήμα 3) που προκαλεί το εκάστοτε συναίσθημα στον ομιλητή ενώ τα ακουστικά προκύπτουν από την επεξεργασία των ακουστικών σημάτων που παράγει [4]. Η βιντεοσκόπηση αυτή, που αποτελεί το σύνολο των δεδομένων μας έχει χωριστεί σε διαδοχικά (ως προς τον χρόνο) τμήματα τα οποία ονομάζονται tunes και το καθένα από αυτά αποτελείται από έναν αριθμό από καρτέ (frames). Κάθε frame από μόνο του αντιστοιχεί σε μία κλάση. Τα frames που ανήκουν στο ίδιο tune ανήκουν προφανώς στην ίδια κλάση και αυτή είναι και η κλάση του tune. Έχουμε, λοιπόν, 477 tunes (το κάθε tune έχει έναν αριθμό από frames) στα οποία υπάρχουν 49 εισοδοί και 5 έξοδοι. Από τις 49 εισόδους, οι 17 αντιστοιχούν στα χαρακτηριστικά του προσώπου και οι 32 στην ομιλία. Συνοπτικά, τα στοιχεία δίνονται στον πίνακα 1:



Σχήμα 3: Εκφράσεις Προσώπου Που Αντιπροσωπεύουν Συναισθήματα

Tunes	είσοδοι	πρόσωπο	φωνητικά	έξοδοι
477	49	17	32	5

Πινάκας 1: Στοιχεία Συνόλου Δεδομένων

Στην πειραματική αυτή διαδικασία ασχολούμαστε με 17 παραμέτρους οι οποίες αντιπροσωπεύουν “καταστάσεις” στο πρόσωπο και ονομάζονται Facial Animation Parameters–FAPs (Παράμετροι Κίνησης Προσώπου). Το μέγεθος της μετακίνησης που περιγράφεται από ένα FAP εκφράζεται μέσω κάποιων μονάδων μέτρησης που ονομάζονται Facial Animation Parameter Units (FAPU), που αντιπροσωπεύουν τμήματα συγκεκριμένων αποστάσεων στο πρόσωπο. Οι περιστροφές περιγράφονται σαν τμήματα ενός ακτινίου.

Τα ακουστικά χαρακτηριστικά βασίζονται αποκλειστικά στην προσωδία και σχετίζονται με τον τόνο και τον ρυθμό ενώ το σύνολο των φωνητικών δεδομένων είναι χωρισμένο σε τμήματα. Κάθε tune αναλύθηκε με μία μέθοδο που χρησιμοποιεί την προσωδική αναπαράσταση βασισμένη στη μέθοδο που ονομάζεται ‘Prosogram’. Η Prosogram βασίζεται στην επεξεργασία των χαρακτηριστικών της βασικής συχνότητας των φωνητικών πυρήνων (το ελάχιστο από τα τμήματα στα οποία μπορεί να “διασπαστεί” η ομιλία). Δίνει δηλαδή σε κάθε φωνητικό πυρήνα έναν τόνο και ένα μήκος.

Στη μελέτη αυτή, το μεγάλο πλήθος συναισθημάτων που μπορεί να νιώσει ο άνθρωπος οργανώνεται σε 5 γενικές κατηγορίες συναισθηματικής κατάστασης που ονομάζονται κλάσεις.

Οι πέντε κλάσεις αντιπροσωπεύουν τα τέσσερα τεταρτημόρια συν την αρχή των αξόνων. Αυτή είναι μια κωδικοποίηση για τις εξής γενικές κατηγορίες συναισθηματικής κατάστασης:

- Θετική ενεργητική (τεταρτημόριο 1)
- Θετική παθητική (τεταρτημόριο 2)
- Αρνητική ενεργητική (τεταρτημόριο 3)

- Αρνητική παθητική (τεταρτημόριο 4)
- Ουδέτερη (αρχή των αξόνων)

Σε ένα σύστημα σαν το δικό μας, με τόσο μεγάλο πλήθος εισόδων, θα περίμενε κανείς πολύ μεγάλο πλήθος κρυφών κόμβων οι οποίοι είναι εκθετικά ανάλογοι των εισόδων. Αυτό, βέβαια, θα είχε ως αποτέλεσμα και την εκθετική αύξηση της πολυπλοκότητάς του. Στόχος μας ήταν η αποφυγή αυτού του φαινομένου διατηρώντας, όμως, την καλή απόδοση του δικτύου.

Η ανάλυση ευαισθησίας, εφαρμοζόμενη στο νευρωνικό μας, μας υποδείκνυε σε κάθε εκπαίδευση τις 12 εισόδους με τη μικρότερη επιρροή στο τελικό αποτέλεσμα. Επαναδοκιμάζοντας το δίκτυο χωρίς αυτές, παρατηρήσαμε βελτίωση της απόδοσης με ποσοστό 3 % κατά μέσο όρο. Καταλήξαμε, λοιπόν, σε έναν μικρότερο όγκο δεδομένων (37 εισοδοί) και άρα ένα γρηγορότερο σύστημα του οποίου η απόδοση είναι ακόμα καλύτερη.

Τονίζουμε ότι στην παρούσα εργασία επιδιώξαμε τη μέγιστη μείωση εισόδων η οποία παράλληλα να επιφέρει και αύξηση στην απόδοση. Εναλλακτικά, θα μπορούσαμε να στοχεύσουμε πρωτίστως στην αύξηση της απόδοσης, ώστε να επιτύχουμε μεγαλύτερα ποσοστά, και επακόλουθα να είχαμε μείωση των εισόδων, αλλά σε μικρότερο βαθμό.

5. Συμπεράσματα & Μελλοντικές επεκτάσεις

Έχοντας στα χέρια μας τα χαρακτηριστικά μιας μαγνητοσκόπησης και τις γενικές θεωρητικές γνώσεις για τα αναδρομικά νευρωνικά δίκτυα και την ανάλυση ευαισθησίας προσπαθήσαμε να συμβάλουμε με τα αποτελέσματά μας στο πολύ ευρύ επιστημονικό πεδίο της αναγνώρισης των συναισθημάτων.

Τα προβλήματα που έπρεπε να αντιμετωπιστούν εμφανίστηκαν κιάλας κατά την επιλογή των tunes που θα χρησιμοποιούνταν καθώς και κατά τον διαχωρισμό των συνόλων εκπαίδευσης και δοκιμής λόγω της φύσης των δεδομένων. Το μεγάλο πλήθος δοκιμών και πειραμάτων μας οδήγησε στην κατάλληλα επιλογή των διαφόρων παραμέτρων του προβλήματος ώστε να καταλήξουμε τελικά σε ένα καλά εκπαιδευμένο αναδρομικό νευρωνικό δίκτυο.

Η επόμενη πρόκληση αφορούσε στην εφαρμογή της ανάλυσης ευαισθησίας σε ανατροφοδοτούμενο δίκτυο. Η προσπάθεια αυτή περιλάμβανε το να ληφθεί υπ' όψιν η ιδιαιτερότητα των αναδρομικών δικτύων που είναι η μνήμη και να ενσωματωθεί αυτή στον αλγόριθμο ενός στατιστικού εργαλείου όπως είναι η ανάλυση ευαισθησίας.

Κάνοντας την “διαδρομή” αυτή συνειδητοποιήσαμε την δύναμη των αναδρομικών δικτύων αλλά και τις δυσκολίες που αυτή συνεπάγεται. Είδαμε σε βάθος την φιλοσοφία της λειτουργίας και της εκπαίδευσής τους και

δημιουργήσαμε ένα καλά εκπαιδευμένο δίκτυο που να επεξεργάζεται χαρακτηριστικά προσώπου και ομιλίας και να αναγνωρίζει την συναισθηματική κατάσταση.

Επίσης αποδείξαμε τη μεγάλη σημασία της ανάλυσης ευαισθησίας για την εξαγωγή πολύ χρήσιμων πληροφοριών για τις εισόδους του συστήματος οι οποίες μας οδήγησαν στην βελτιστοποίησή του. Πιο συγκεκριμένα, η μέθοδος αυτή μας υπέδειξε τις εισόδους με την μικρότερη επιρροή στην λήψη απόφασης του δικτύου. Γνωρίζοντας τις εισόδους αυτές κάναμε επαναδοκιμή του δικτύου, χωρίς αυτές λαμβάνοντας, έτσι, ένα λιγότερο πολύπλοκο και πιο αποδοτικό σύστημα. Η σημασία της μείωσης της πολυπλοκότητας φαίνεται πολύ καθαρά αν αναλογιστούμε ότι όσο αυξάνονται οι εισοδοί ενός νευρωνικού, το πλήθος των κρυφών κόμβων που απαιτείται αυξάνεται εκθετικά.

Εξ' ίσου σημαντικό επακόλουθο της εφαρμογής της ανάλυσης ευαισθησίας στο δίκτυό μας είναι και ο καθορισμός των χαρακτηριστικών που ασκούν τη μεγαλύτερη επιρροή σε κάθε έξοδο δίνοντάς μας μία πιο συγκεκριμένη εικόνα για το ποιά είναι τα χαρακτηριστικά στα οποία πρέπει να εστιάσουμε όταν θέλουμε να ανιχνεύσουμε ένα συγκεκριμένο συναίσθημα.

Η περαιτέρω διερεύνηση του αντικειμένου της εργασίας αυτής θα μπορούσε να περιλαμβάνει μεγαλύτερο πλήθος πειραματικών δεδομένων για την επιβεβαίωση της αποδοτικότητάς της και σε άλλα προβλήματα.

Μια ακόμα προσέγγιση θα μπορούσε να περιλαμβάνει μια πρώτη μείωση των εισόδων, ίσως με βάση την μεταξύ τους συσχέτιση, και ακολούθως την επιπλέον μείωση που προκύπτει από την ανάλυση ευαισθησίας.

Πρωταρχικός στόχος ήταν η μείωση των εισόδων σε όσο το δυνατόν μεγαλύτερο βαθμό, έχοντας επιπλέον βελτίωση στην απόδοση. Αν ο στόχος είναι, κατά κύριο λόγο, η αύξηση της απόδοσης χωρίς να στοχεύουμε στις ελάχιστες δυνατές εισόδους, μπορούμε να αφαιρέσουμε μικρότερο αριθμό εισόδων και να πετύχουμε ακόμα υψηλότερες αποδόσεις

6. Αναφορές

- [1] Addison, D.; Wermter, S. & Arevian, G. (2003), 'A comparison of feature extraction and selection techniques *Proceedings of International Conference on Artificial Neural Networks*, 212--215.
- [2] Bishop, C. (2006), *Pattern recognition and machine learning*, Springer New York..
- [3] Boden, M. (1999), 'A guide to recurrent neural networks and backpropagation', *The DALLAS project. Report from the NUTEK-supported project AIS-8: Application of Data Analysis with Learning Systems*.
- [4] Caridakis, G.; Malatesta, L.; Kessous, L.; Amir, N.; Raouzaoui, A. & Karpouzis, K. (2006), 'Modeling

naturalistic affective states via facial and vocal expressions recognition"Proceedings of the 8th international conference on Multimodal interfaces', ACM New York, NY, USA, 146--154.

- [5] Engelbrecht, A.; Cloete, I. & Zurada, J. (1995), 'Determining the significance of input parameters using sensitivity analysis'Proceedings of the International Workshop on Artificial Neural Networks: From Natural to Artificial Neural Computation', Springer-Verlag London, UK, 382--388.
- [6] Guijarro-Berdinas, B.; Fontenla-Romero, O.; Perez-

Sanchez, B. & Alonso-Betanzos, A. (2006), 'A new initialization method for neural networks using sensitivity analysis', *Proceedings of the 2006 International Conference on Mathematical and Statistical Modeling*.

- [7] Picard, R.; Vyzas, E. & Healey, J. (2001), 'Toward machine emotional intelligence: analysis of affectivephysiological state', *IEEE transactions on pattern analysis and machine intelligence* **23**(10), 1175--1191.