

**ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
Δ.Π.Μ.Σ. ΣΥΣΤΗΜΑΤΑ ΑΥΤΟΜΑΤΙΣΜΟΥ**

**ΜΕΤΑΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ**

**«ΧΡΗΣΗ ΤΕΧΝΗΤΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ ΓΙΑ ΤΗΝ ΠΡΟΒΛΕΨΗ  
ΕΠΙΦΑΝΕΙΑΚΗΣ ΤΡΑΧΥΤΗΤΑΣ ΣΤΗΝ ΚΑΤΕΡΓΑΣΙΑ ΤΗΣ ΤΟΡΝΕΥΣΗΣ  
ΥΛΙΚΩΝ ΜΕ ΥΨΗΛΗ ΣΚΛΗΡΟΤΗΤΑ»**

**ΛΩΛΟΣ ΕΜΜΑΝΟΥΗΛ**

**A.M. 02114112**

**ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ:**

**Δ. Ε. ΜΑΝΩΛΑΚΟΣ**

**ΑΘΗΝΑ 2016**

*Θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή μου κ.Μανωλάκο Δημήτριο για την ανάθεση του θέματος και την εμπιστοσύνη του σε εμένα, καθώς και τους κ.Μαρκόπουλο Άγγελο και κ.Κάρκαλο Νικόλαο για όλη την καθοδήγηση και τη βοήθεια που μου παρείχαν κατά την εκπόνηση της παρούσας μεταπτυχιακής εργασίας.*

# ΠΕΡΙΕΧΟΜΕΝΑ

## ΠΕΡΙΛΗΨΗ

### Α' ΘΕΩΡΗΤΙΚΟ ΜΕΡΟΣ

#### ΚΕΦΑΛΑΙΟ 1<sup>ο</sup> : Η ΚΑΤΕΡΓΑΣΙΑ ΤΗΣ ΤΟΡΝΕΥΣΗΣ

1.1	Η Κατεργασία της τórνευσης.....σελ.8
1.2	Είδη τórνευσης.....σελ.9
1.3	Συνθήκες κοπής στην κατεργασία της τórνευσης.....σελ.12
1.4	Δυνάμεις στην κοπή.....σελ.12
1.5	Υπολογισμός της κύριας συνιστώσας της δύναμης κοπής.....σελ.13
1.6	Ρυθμός αφαίρεσης υλικού.....σελ.16
1.7	Τα κοπτικά εργαλεία στην κατεργασία της τórνευσης.....σελ.17
1.8	Τα υγρά κοπής στην κατεργασία της κοπής.....σελ.20
1.9	Επιφανειακή τραχύτητα έπειτα από τórνευση.....σελ.22
1.10	Τórνευση σκληρών υλικών (hard turning).....σελ.24

#### ΚΕΦΑΛΑΙΟ 2<sup>ο</sup> : ΤΑ ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

2.1	Η ιστορία των τεχνητών νευρωνικών δικτύων.....σελ.26
2.2	Τι είναι τα τεχνητά νευρωνικά δίκτυα.....σελ.26
2.3	Δυνατότητες και πλεονεκτήματα των νευρωνικών δικτύων.....σελ.27
2.4	Εφαρμογές των τεχνητών νευρωνικών δικτύων.....σελ.28
2.5	Λειτουργία του βιολογικού νευρωνικού δικτύου.....σελ.28
2.6	Το μοντέλο του τεχνητού νευρώνα.....σελ.30
2.7	Τα τεχνητά νευρωνικά δίκτυα (ΤΝΔ).....σελ.35
2.8	Αρχιτεκτονικές τεχνητών νευρωνικών δικτύων.....σελ.35
2.9	Διαδικασίες εκπαίδευσης ΤΝΔ.....σελ.39
2.10	Εργασίες μάθησης.....σελ.44
2.11	Το Perceptron.....σελ.49
2.12	Ο αλγόριθμος ελαχίστων μέσων τετραγώνων για τη μέθοδο διόρθωσης σφάλματος.....σελ.52
2.13	Το Perceptron πολλαπλών επιπέδων.....σελ.58

#### ΚΕΦΑΛΑΙΟ 3<sup>ο</sup> : ΧΡΗΣΗ ΤΝΔ ΓΙΑ ΤΗ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ ΚΑΤΕΡΓΑΣΙΩΝ

3.1	Εισαγωγή.....σελ.69
3.2	Εργασίες στη χρήση των ΤΝΔ για τη βελτιστοποίηση κατεργασιών.....σελ.69

## **Β' ΠΕΙΡΑΜΑΤΙΚΟ ΜΕΡΟΣ**

### **ΚΕΦΑΛΑΙΟ 4<sup>ο</sup> : ΠΕΙΡΑΜΑΤΙΚΟ ΜΕΡΟΣ**

4.1 Εισαγωγή.....σελ.79
4.2 Περιγραφή της εργασίας προς σύγκριση.....σελ.79
4.3 Τεχνητά νευρωνικά δίκτυα.....σελ.84
4.4 Συμπεράσματα.....σελ.133

### **ΠΑΡΑΡΤΗΜΑ**

### **ΒΙΒΛΙΟΓΡΑΦΙΑ**

## ΠΕΡΙΛΗΨΗ

Στα πλαίσια της παρούσας μεταπτυχιακής εργασίας επιχειρήσαμε να βρούμε το βέλτιστο Τεχνητό Νευρωνικό Δίκτυο (ΤΝΔ), το οποίο θα μπορούσε να κάνει πρόβλεψη της τιμής της επιφανειακής τραχύτητας έπειτα από τόννευση του χάλυβα με υψηλή σκληρότητα AISI 4340 (69HRC). Αυτό έγινε καθώς προσπαθήσαμε να συγκρίνουμε την απόδοση των τεχνητών νευρωνικών δικτύων με την απόδοση των μοντέλων που προτείνουν στην δημοσιευμένη εργασία τους οι A.Agrawal, S.Goel, W.Bin Rashid και M.Price με τίτλο " Prediction of surface roughness during hard turning of AISI 4340 steel (69 HRC)" [1] . Στα μοντέλα αυτά χρησιμοποιήθηκαν πειραματικά δεδομένα τα οποία αποτέλεσαν και δεδομένα εκπαίδευσης των νευρωνικών δικτύων που αναπτύξαμε εμείς. Στη δημοσιευμένη αυτή εργασία αναφέρεται ότι ένα από τα τρία μοντέλα (το μοντέλο παλινδρομησης με τυχαία δάση) που αναπτύχθηκαν-εξετάστηκαν, έχει αποδειχθεί να είναι ανώτερο πολλών άλλων υπολογιστικών μεθόδων μια από τις οποίες είναι και τα νευρωνικά δίκτυα.

Τα νευρωνικά δίκτυα που αναπτύξαμε ήταν πρόσθιας τροφοδότησης (multi-layer perceptron - MLP) και προγραμματίστηκαν στη MATLAB. Για να καταλήξουμε σε ένα τέτοιο ΤΝΔ δοκιμάσαμε πολλούς διαφορετικούς συνδυασμούς αριθμών νευρώνων για δίκτυα με δύο κρυφά επίπεδα. Για κάθε αρχιτεκτονική, όπου εκπαίδευσαν 5000 ΤΝΔ, φροντίσαμε τελικά να φιλτράρουμε τα εκπαιδευμένα αυτά δίκτυα εντοπίζοντας και αποθηκεύοντας εκείνο το οποίο είχε το μικρότερο σφάλμα ελέγχου. Έπειτα, για την επιλογή της βέλτιστης αρχιτεκτονικής εντοπίσαμε αρχικά τις δομές με τα λιγότερα δεδομένα που είχαν σφάλμα >10% της ζητούμενης επιφανειακής τραχύτητας και στη συνέχεια προχωρήσαμε σε περαιτέρω διερεύνηση βασισμένη στη σύγκριση των μέσων τετραγωνικών σφαλμάτων (δίνοντας μεγαλύτερη βαρύτητα στο μέσο τετραγωνικό σφάλμα ελέγχου) μεταξύ των δομών αυτών.

Για την εκπαίδευση των δικτύων χωρίσαμε τα δεδομένα σε τρία υποσύνολα όπου επιλέχθηκε ένας διαχωρισμός 70%-15%-15%. Το πρώτο υποσύνολο χρησιμοποιήθηκε για την εκπαίδευση, το δεύτερο για τον τερματισμό της εκπαίδευσης και την αποφυγή υπερεκπαίδευσης (κριτήριο του validation set) και το τρίτο για τον έλεγχο τους από εμάς και την επιλογή βέλτιστου δικτύου σε κάθε διαφορετική δομή. Όλα τα παραπάνω επαναλήφθηκαν για τρεις διαφορετικές μεθόδους-αλγόριθμους εκπαίδευσης: τον προσαρμοστικό αλγόριθμο ανάστροφης διάδοσης της πιο απότομης κατάβασης με χρήση όρου ορμής (gradient descent with momentum) "**trainGDX**", τον αλγόριθμο ανάστροφης διάδοσης των Levenberg-Marquardt "**trainLM**" και τον αλγόριθμο συζυγών κλίσεων scaled conjugate gradient "**trainSCG**".

Για την καλύτερη σύγκλιση κατά την εκπαίδευση των ΤΝΔ, στα δεδομένα εκπαίδευσης και ελέγχου του δικτύου εφαρμόστηκε η τεχνική της κανονικοποίησης δεδομένων σε ένα εύρος μεταξύ 0-1 διαιρώντας, για κάθε παράμετρο, όλες τις τιμές με την μέγιστη τιμή της παραμέτρου, αντίστοιχα.

Όσον αφορά το μέσο τετραγωνικό σφάλμα εκπαίδευσης, φαίνεται η μέθοδος trainGDX να έχει δώσει τα χειρότερα αποτελέσματα από τις τρεις μεθόδους. Ανάμεσα στην trainLM και την trainSCG φαίνεται η μέθοδος trainLM να έχει οριακά μικρότερα σφάλματα. Όσον αφορά το μέσο τετραγωνικό σφάλμα ελέγχου, φαίνεται η μέθοδος trainGDX να έχει οριακά χειρότερα αποτελέσματα από τη μέθοδο trainLM, ενώ μεταξύ των μεθόδων trainLM και trainSCG διαπιστώνουμε και ότι η μέθοδος trainSCG έδωσε μικρότερα σφάλματα ελέγχου από τη μέθοδο trainLM.

Όσον αφορά τον ρόλο που παίζει το μέγεθος του δικτύου (ο αριθμός των νευρώνων στα κρυφά επίπεδα) στην απόδοσή του, μπορούμε να παρατηρήσουμε ότι καθώς κινούμαστε προς την αύξηση του αριθμού νευρώνων στο πρώτο κρυφό επίπεδο, φαίνεται να αυξάνεται ελαφρώς το σφάλμα ελέγχου για όλες τις μεθόδους. Για το σφάλμα εκπαίδευσης δεν παρατηρείται κάποιο αντίστοιχο συμπέρασμα.

Από τα τρία βέλτιστα ΤΝΔ για κάθε μέθοδο εκπαίδευσης επιλέξαμε για σύγκριση με το μοντέλο παλινδρόμησης με τυχαία δάση, εκείνο το οποίο είχε το μικρότερο μέσο τετραγωνικό σφάλμα εκπαίδευσης και ελέγχου. Το δίκτυο αυτό είναι το βέλτιστο ΤΝΔ της μεθόδου trainLM με δομή 3-11-5-1. Το ΤΝΔ που επιλέξαμε ως βέλτιστο δίνει σαφώς καλύτερα αποτελέσματα από το μοντέλο παλινδρόμησης με τυχαία δάση το οποίο, από τα συνολικά 39 πειράματα, δίνει οριακά καλύτερο αποτέλεσμα μόνο σε 10 πειράματα ενώ παρουσιάζει σημαντικά μεγαλύτερο σφάλμα σε σχέση με τα αποτελέσματα του ΤΝΔ σε 8 δεδομένα.

## **ABSTRACT**

In this study we tried to find the optimum Artificial Neural Network (ANN) which would be able to predict the surface roughness of the high hardness AISI 4340 steel (69HRC) after turning. For this purpose we tried to compare the performance of artificial neural networks to the performance of the models which are suggested from A.Agrawal, S.Goel, W.Bin Rashid and M.Price in their paper with title "Prediction of surface roughness during hard turning of AISI 4340 steel (69 HRC)" [1]. The experimental data that were used in these models, became the training data of the artificial neural networks that we developed. In the paper it is mentioned that one of these three models (the random forest regression model - RFR) that were developed-tested, is proved to be superior among other numerical models, one of them being the artificial neural networks.

The artificial neural networks that we created were feed forward networks (multi-layer perceptron - MLP) and they were developed in MATLAB. In order to end up to such an ANN we tried many different combinations of the number of neurons for two hidden layer networks. For each architecture, where we trained 5000 ANNs, we managed to filter these trained networks by locating and saving only the network with the lowest test error. Afterwards, for the choice of the best architecture, we firstly located the structures which had the least number of data with error >10% for the desired surface roughness value and consequently we proceeded to further investigation among these structures, based to the comparison of the mean squared errors (by attaching greater importance to the mean square error of the test set).

For the train of our networks we divided the data to three sets of 70%-15%-15%. The first set was used for training, the second for the termination of training and the avoidance of overfitting (validation set criterion) and the third set for testing the networks and choosing the best ANN for each and every structure. All the above were repeated for three different training methods-algorithms: the gradient descent with momentum and adaptive learning rate backpropagation "trainGDX", the Levenberg-Marquardt backpropagation "trainLM" and the scaled conjugate gradient backpropagation "trainSCG".

For better convergence during the training of the ANNs, we applied to the training data the method of data normalization to a range between 0-1 by dividing, for each parameter, all the values with the maximum value of the parameter, respectively.

As far as mean square error of the training set is concerned, it is obvious from the results that the trainGDX algorithm gives the worst prediction values among the three methods. Between trainLM and trainSCG algorithms, we can conclude that trainLM gives us slightly smaller training error. As far as mean square error of the test set is concerned, we can see from the results that trainGDX has marginally worse prediction values than trainLM, while between trainLM and trainSCG algorithms, trainSCG gives us lower test error.

As regards the role of the size of the ANN (number of neurons for each hidden layer) to its performance, we can conclude that as we augment the number of neurons for the first hidden layer, we can observe an increasment of the test error for all the three algorithms. For the training error such an observation is not visible.

Among the three optimal artificial neural networks (one for each algorithm), we chose the ANN with the lower mean square error of the training and the test set in order to compare it with the random forest regression model. This network was the optimal ANN of the trainLM algorithm with structure 3-11-5-1. This specific network gives clearly better results than the RFR model which, from the total 39 experiments, gives us barely better predictions for only 10 experiments while it has significantly higher error than this specific ANN for 8 experiments.

**Α' ΘΕΩΡΗΤΙΚΟ ΜΕΡΟΣ**

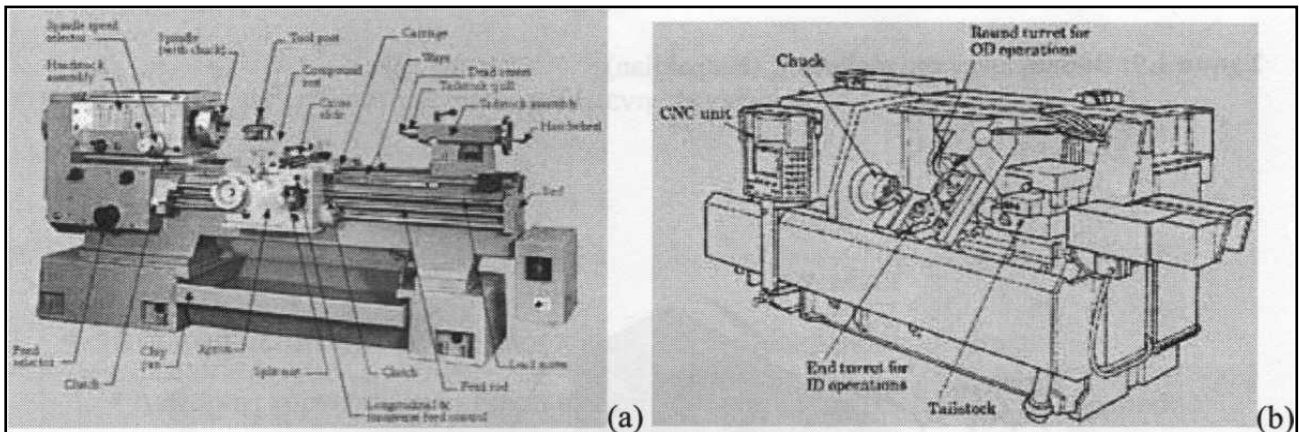
## ΚΕΦΑΛΑΙΟ 1<sup>ο</sup> Η ΤΟΡΝΕΥΣΗ

### 1.1 Η κατεργασία της τórνευσης

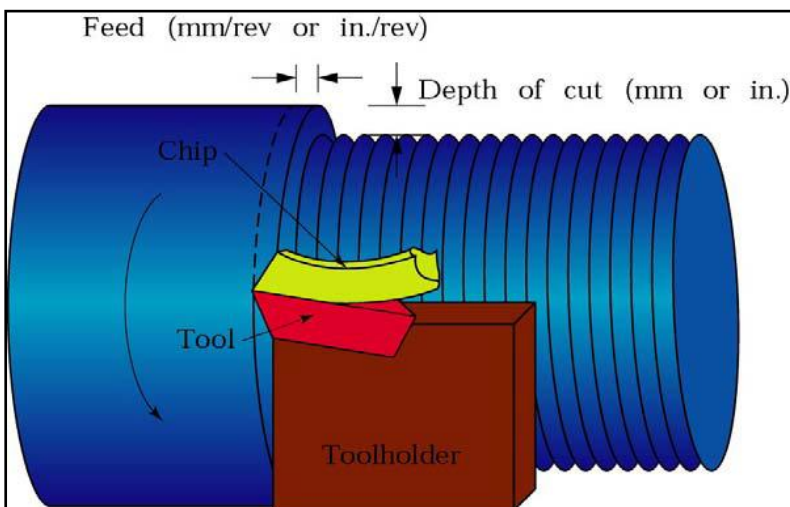
Η τórνευση ανήκει στις κατεργασίες κοπής με κύρια κίνηση την περιστροφική κίνηση της ατράκτου, όπου προσδένεται κατάλληλα το προς κατεργασία τεμάχιο, και δευτερεύουσα τη μεταφορική κίνηση της πρόωσης (κίνηση του κοπτικού εργαλείου), εκτελείται δε με κοπτικό εργαλείο καθορισμένης γεωμετρίας και απλής σημειακής επαφής. Χαρακτηριστικός τύπος συμβατικού τórνου όπως και η σύγχρονη μετεξέλιξη του φαίνονται στο σχήμα 1.1

Όταν η τórνευση πραγματοποιείται σε τεμάχια το υλικό των οποίων έχει σκληρότητα μικρότερη από 45 HRC, αναφερόμαστε στη συμβατική ή κανονική τórνευση (soft turning). Όταν πραγματοποιείται σε κομμάτια που έχουν τιμές σκληρότητας μεγαλύτερες των 45 HRC, το κοπτικό εργαλείο πρέπει να έχει ιδιότητες που να επιτρέπουν την κατεργασία και αναφερόμαστε σε αυτό το είδος τórνευσης ως hard turning.

Σχηματική παράσταση της κατεργασίας και τυποποιημένη μορφή του κοπτικού εργαλείου τórνευσης παρουσιάζονται στα σχήματα 1.2 και 1.3 αντίστοιχα.

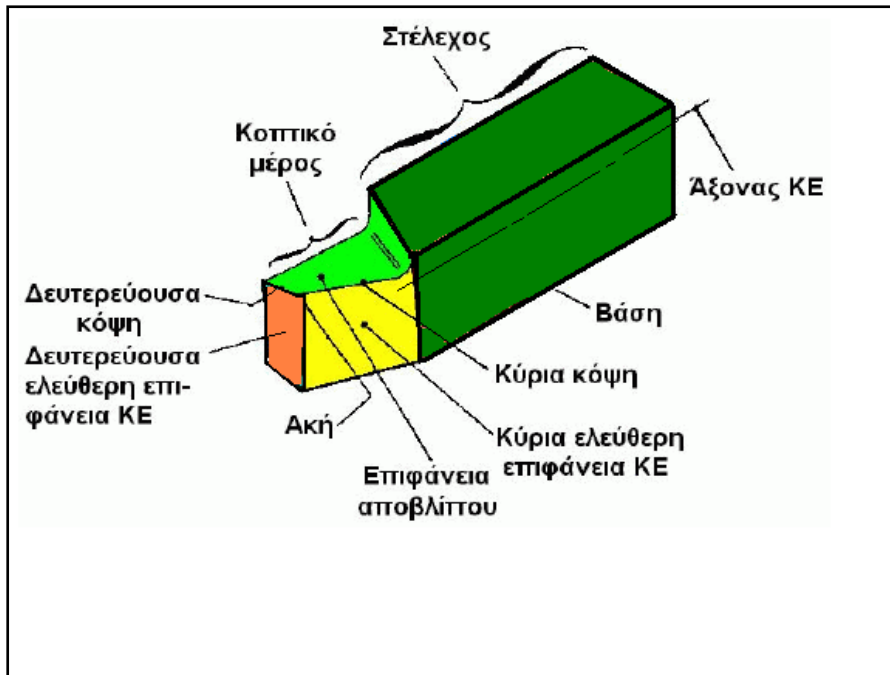


**Σχήμα 1.1** (α) Συμβατικός τórνος και (β) Μηχανή τórνευσης CNC (Kalpakjian) [2].



**Σχήμα 1.2** Βασική αρχή τórνευσης [2].



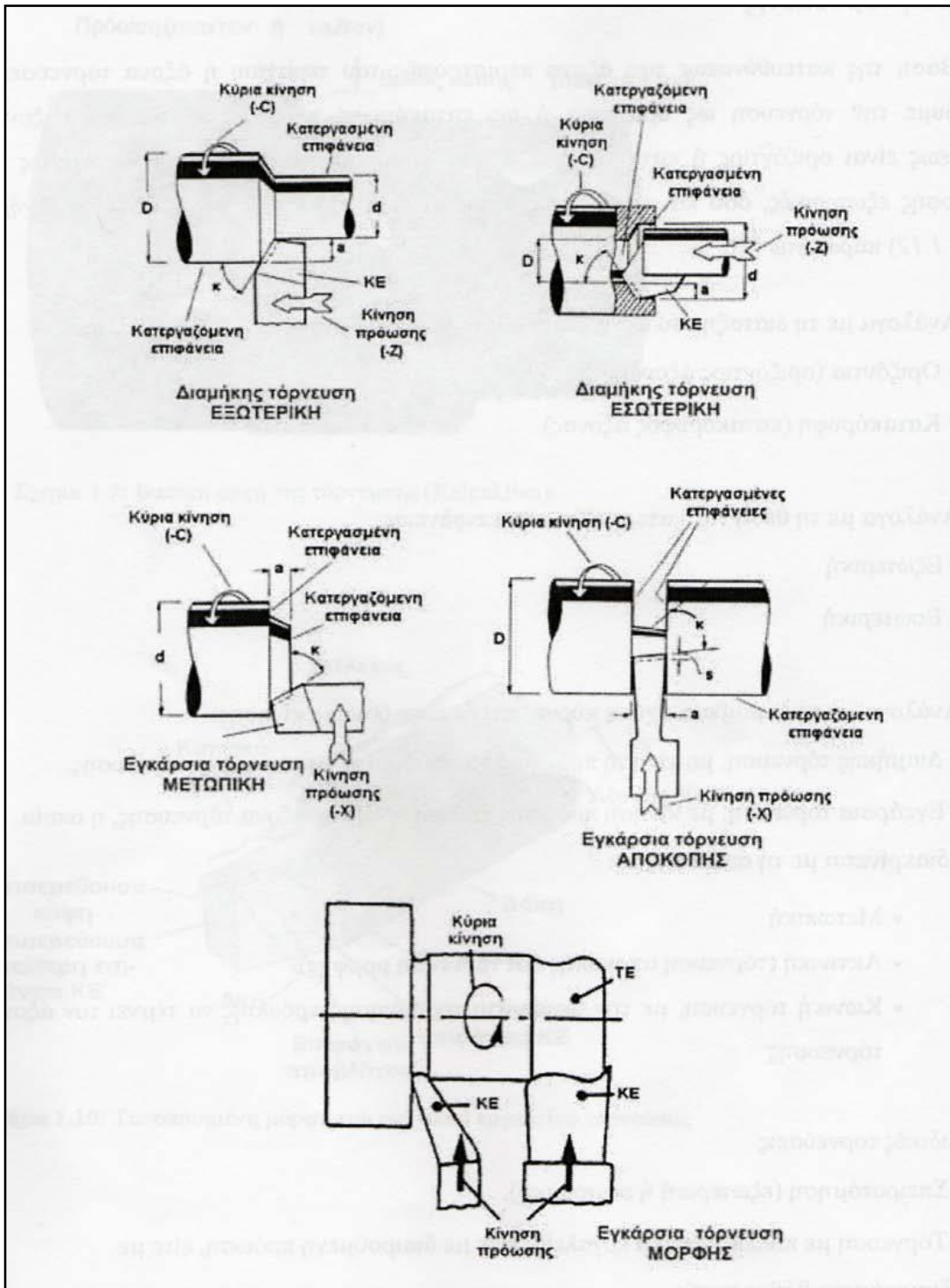


Σχήμα 1.3 Τυποποιημένη μορφή του κοπτικού εργαλείου τόννευσης [2].

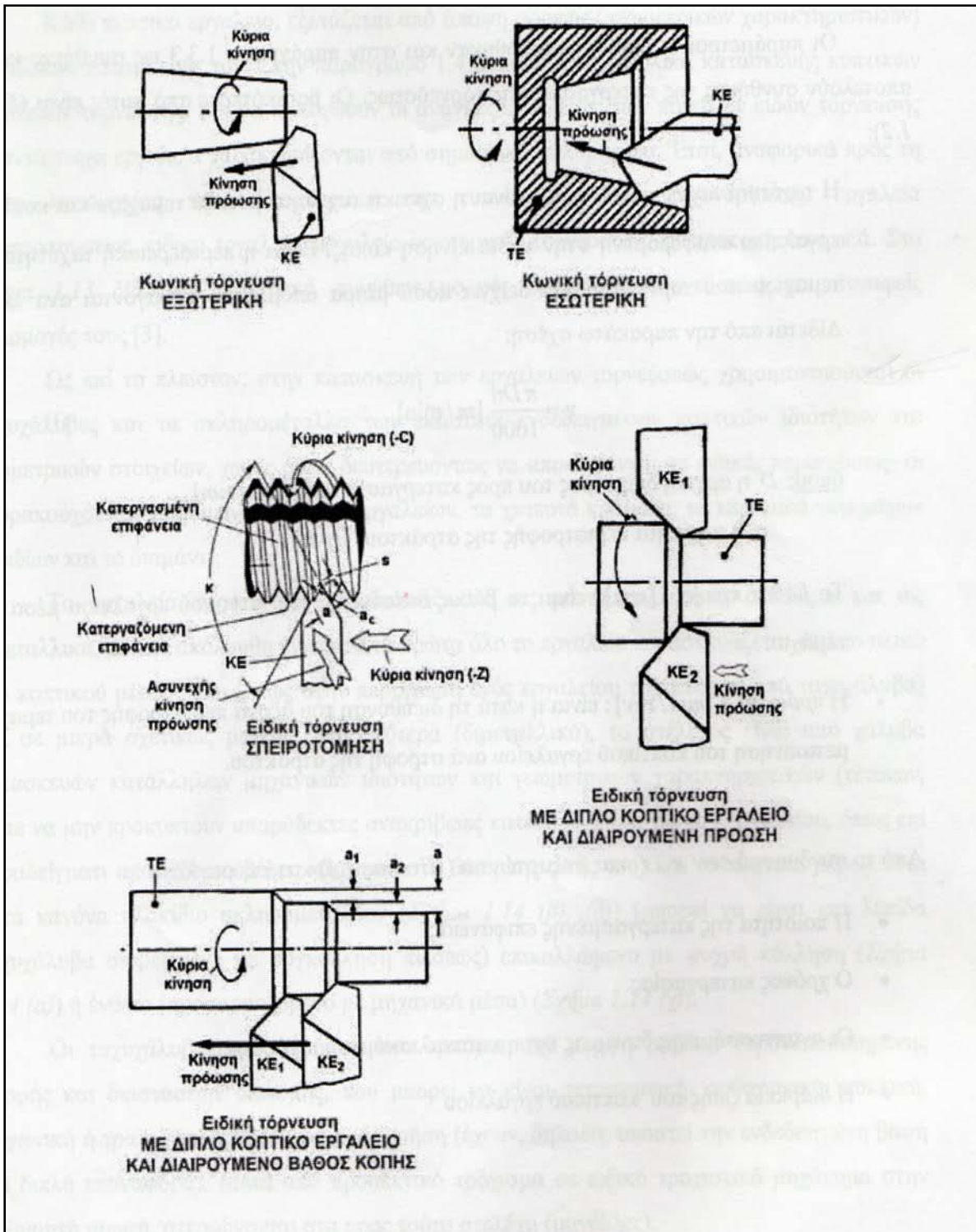
## 1.2 Είδη τόννευσης

Βάσει του άξονα περιστροφής του τεμαχίου ή άξονα τόννευσης διακρίνουμε την τόννευση ως οριζόντια ή ως κατακόρυφη. Γενικά, μπορούν να τόννευτούν επιφάνειες εκ περιστροφής εξωτερικές όσο και εσωτερικές. Έτσι τα είδη τόννευσης ταξινομούνται (σχήμα 1.4 και 1.5) όπως παρουσιάζονται παρακάτω [2,3]:

- Ανάλογα με τη διάταξη του άξονα τόννευσης:
  - ✓ Οριζόντια (οριζόντιος άξονας)
  - ✓ Κατακόρυφη (κατακόρυφος άξονας).
- Ανάλογα με τη θέση της κατεργαζόμενης επιφάνειας:
  - ✓ Εξωτερική
  - ✓ Εσωτερική.
- Ανάλογα με την αμοιβαία σχέση κύριας και δευτερεύουσας κίνησης
  - ✓ Διαμήκης τόννευση, με κίνηση πρόωσης παράλληλη προς τον άξονα τόννευσης.
  - ✓ Εγκάρσια τόννευση, με κίνηση πρόωσης κάθετη προς τον άξονα τόννευσης, η οποία διακρίνεται με τη σειρά της σε:
    - Μετωπική
    - Ακτινική (τόννευση αποκοπής και τόννευση μορφής)
  - ✓ Κωνική τόννευση, με την διεύθυνση της κίνησης πρόωσης να τέμνει τον άξονα τόννευσης
- Ειδικές τόννευσεις:
  - ✓ Σπειροτόμηση (εξωτερική ή εσωτερική),
  - ✓ Τόννευση με πολλά κοπτικά εργαλεία, είτε με διαιρούμενη πρόωση, είτε με διαιρούμενο βάθος κοπής.



Σχήμα 1.4 Διάφορα είδη τόννευσης – Διαμήκης και εγκάρσια τόννευση [2].



Σχήμα 1.5 Διάφορα είδη τόννευσης – Κωνική και ειδικές τονρεύσεις [2].

### 1.3 Συνθήκες κοπής στην κατεργασία της τόννευσης

Οι συνθήκες κοπής αποτελούν τις συνθήκες της κατεργασίας της τόννευσης. Οι βασικότερες από αυτές είναι:

- **Η ταχύτητα κοπής  $v$  [m/min]:** Είναι η σχετική ταχύτητα μεταξύ τεμαχίου και κοπτικού εργαλείου αναφερόμενη στην κύρια κίνηση κοπής. Είναι η περιφερειακή ταχύτητα του τεμαχίου που τορνεύεται και δείχνει πόσα μέτρα αποβλήττου παράγονται ανά λεπτό. Δίδεται από την παρακάτω σχέση:

$$v = \frac{\pi D n}{1000} [m / \text{min}]$$

όπου:

D: η αρχική διάμετρος του προς κατεργασία τεμαχίου [mm]

n: η ταχύτητα περιστροφής της ατράκτου [rpm]

- **Το βάθος κοπής  $a$  [mm]:** Είναι το βάθος διείσδυσης του κοπτικού εργαλείου μέσα στο τεμάχιο.
- **Η πρόωση  $f$  [mm/rev]:** Είναι η μετακίνηση του κοπτικού εργαλείου ανά στροφή της ατράκτου κατά τη διεύθυνση του άξονα περιστροφής του τεμαχίου.

Από το συνδυασμό των  $v$ ,  $f$  και  $a$  εξαρτώνται κατά κύριο λόγο :

- Η ποιότητα της κατεργασμένης επιφάνειας.
- Ο χρόνος κατεργασίας.
- Οι αναπτυσσόμενες δυνάμεις και η καταναλισκόμενη ισχύς κοπής.
- Η διάρκεια ζωής του κοπτικού εργαλείου.

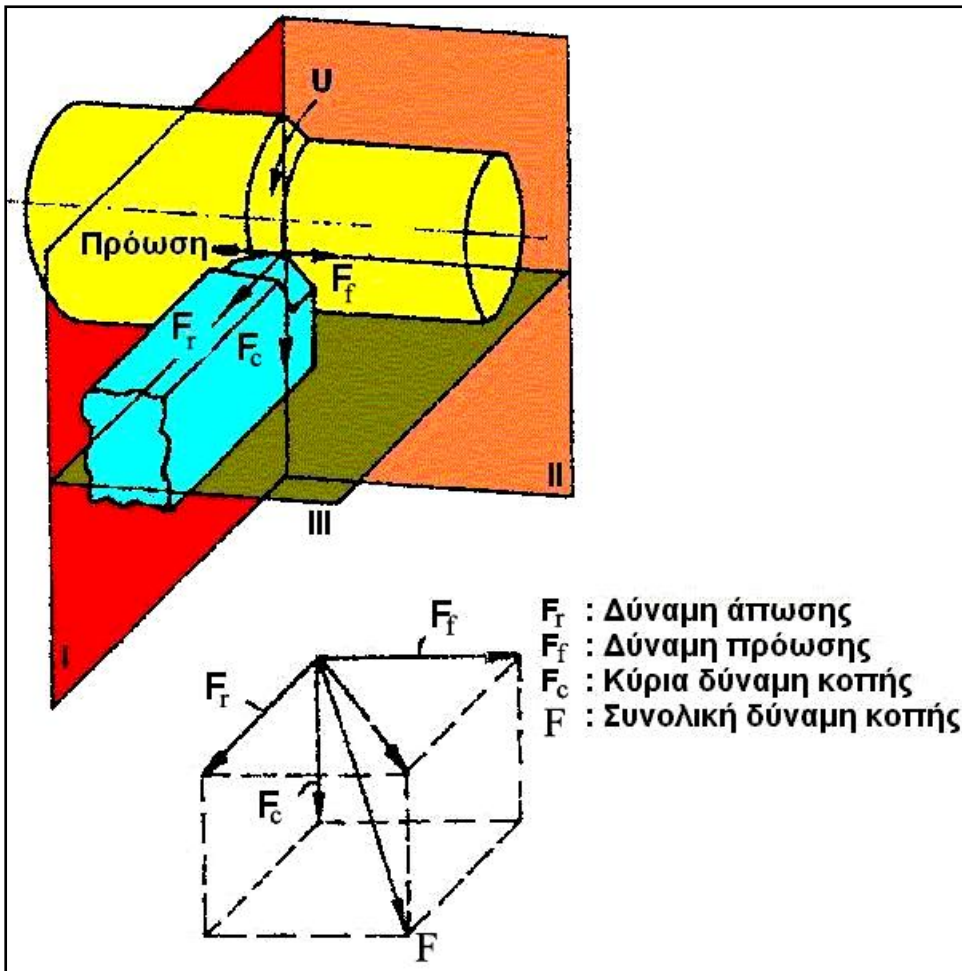
### 1.4 Δυνάμεις στην κοπή

Η συνολική δύναμη κοπής  $F$  αναλύεται στις ακόλουθες συνιστώσες ανά δύο κάθετες μεταξύ τους:

- $F_f$  : Δύναμη πρόωσης ( διεύθυνση X )
- $F_r$  : Δύναμη άπωσης του κοπτικού εργαλείου ( διεύθυνση Y )
- $F_c$  : Κύρια συνιστώσα της δύναμης κοπής ( διεύθυνση Z )

Η δύναμη κοπής  $F_c$  επενεργεί στην άκρη του κοπτικού εργαλείου τείνοντας να το εκτρέψει προς τα κάτω. Αυτή η δύναμη παρέχει την απαιτούμενη ενέργεια κοπής. Η δύναμη πρόωσης  $F_f$  επενεργεί στην διαμήκη κατεύθυνση. Η δύναμη άπωσης  $F_r$  ασκείται στην ακτινική κατεύθυνση και τείνει να απομακρύνει το εργαλείο από το τεμάχιο [2,3].

Προφανώς ισχύει:  $F = \sqrt{F_c^2 + F_r^2 + F_f^2}$



Σχήμα 1.6 Αναπτυσσόμενες δυνάμεις κατά την τórνευση [2].

### 1.5 Υπολογισμός της κύριας συνιστώσας της δύναμης κοπής

Για τον υπολογισμό της κύριας συνιστώσας της δύναμης κοπής μπορούμε να χρησιμοποιήσουμε το "πρότυπο της λοξής κοπής" ή την "ημιεμπειρική μέθοδο Kienzle":

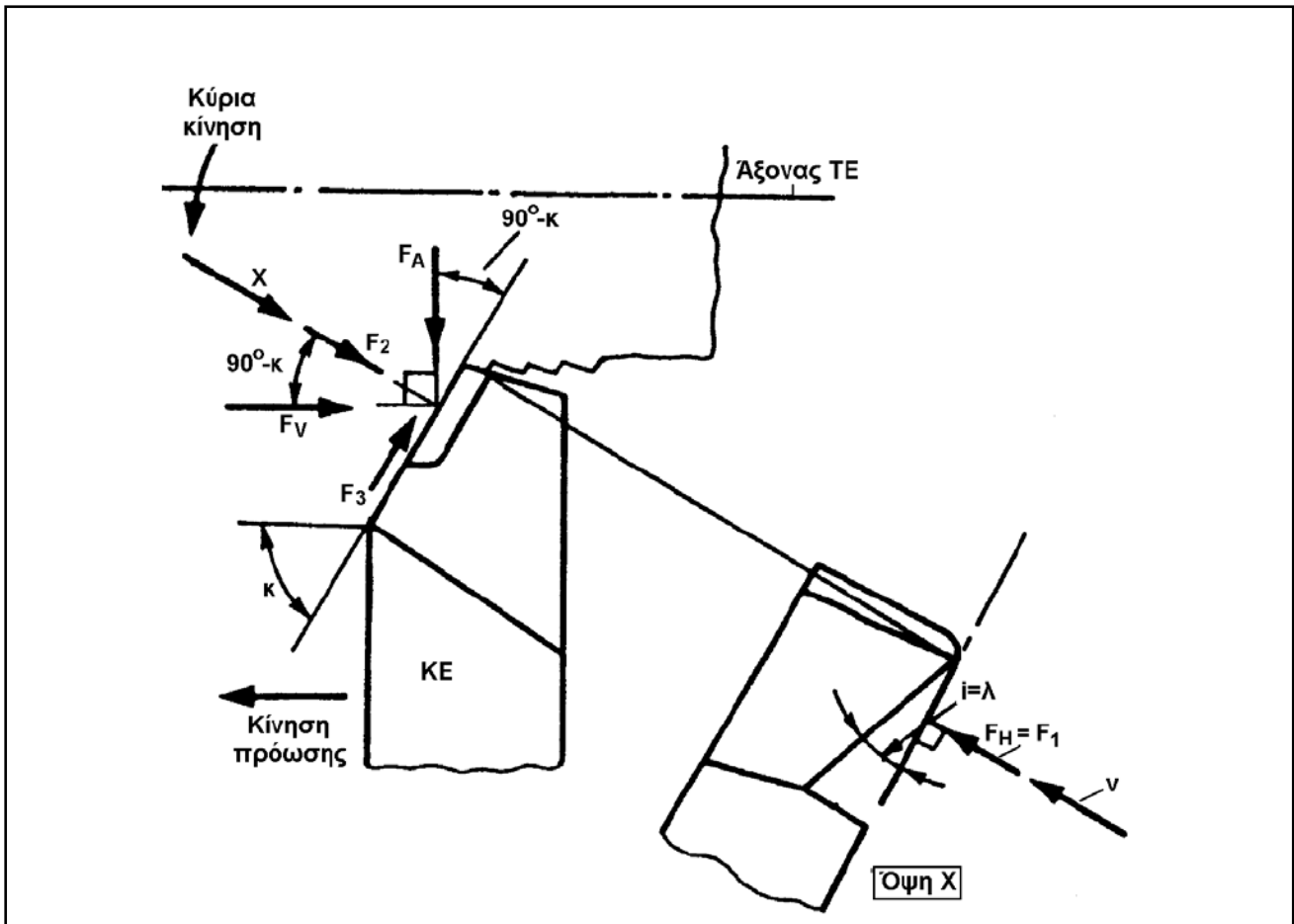
#### 1. Υπολογισμός βάσει του προτύπου της λοξής κοπής:

Κατά την ανάλυση της λοξής κοπής η συνολική δύναμη κοπής  $F$  αναλύθηκε στις συνιστώσες  $F_1$ ,  $F_2$  και  $F_3$ , οι οποίες συνδέονται με τις  $F_f$ ,  $F_r$  και  $F_c$  μέσω των σχέσεων:

a.  $F_c = F_1$

b.  $F_f = F_2 \cdot \sin k + F_3 \cdot \cos k$

c.  $F_r = F_2 \cdot \cos k - F_3 \cdot \sin k$



Σχήμα 1.7 Υπολογισμός της κύριας συνιστώσας της δύναμης κοπής βάσει του προτύπου λοξής κοπής [2].

## 2. Υπολογισμός βάσει της ημιεμπειρικής μεθόδου Kienzle:

Εισάγεται η έννοια της ειδικής αντίστασης κοπής,  $k_s$ , που ορίζεται από τη σχέση:

$$k_s = \frac{F_c}{A} = \frac{F_c}{h \cdot b}$$

Πειραματικά έχει αποδειχθεί ότι, για σταθερή διατομή αποβλίττου, η ειδική αντίσταση κοπής εξαρτάται από το πάχος αποβλίττου σύμφωνα με τη σχέση:

$$k_s = k_1 \cdot h^{-\zeta}$$

όπου:

$k_1$ : Η ειδική αντίσταση κοπής του υλικού τεμαχίου για διατομή αποβλίττου

$A=1\text{mm}^2$  ( $b=h=1\text{mm}$ ) και

$\zeta$ : σταθερά

Άρα, η κύρια συνιστώσα της δύναμης κοπής θα υπολογίζεται από τη σχέση:

$$F_c = k_s \cdot A = k_s \cdot b \cdot h = b \cdot k_1 \cdot h^{(1-\zeta)} \quad (\text{σε daN})$$

Η πειραματική εργασία του Kienzle πραγματοποιήθηκε υπό τις εξής συνθήκες κοπής:

- $v = 90-125 \text{ m/min}$
- $h = 0.06-2.5 \text{ mm}$
- $\gamma = 6^\circ$  για χάλυβα και  $\gamma = \pm 2^\circ$  για χυτοσίδηρο
- $\alpha = 5^\circ$
- Υλικό κοπτικού εργαλείου: Σκληρομέταλλο

Για διαφορετικές συνθήκες κοπής πρέπει να γίνει διόρθωση της τιμής  $F_c$  σύμφωνα με τη σχέση:

$$F_{c\delta} = K_\gamma \cdot K_v \cdot K_\varepsilon \cdot K_\phi \cdot F_H$$

όπου:

- $K_\gamma$ : συντελεστής διόρθωσης λόγω διαφορετικής γωνίας αποβλίττου του κοπτικού εργαλείου, η οποία κυμαίνεται μεταξύ των τιμών  $-20^\circ$  και  $+30^\circ$ , σύμφωνα με τη σχέση:

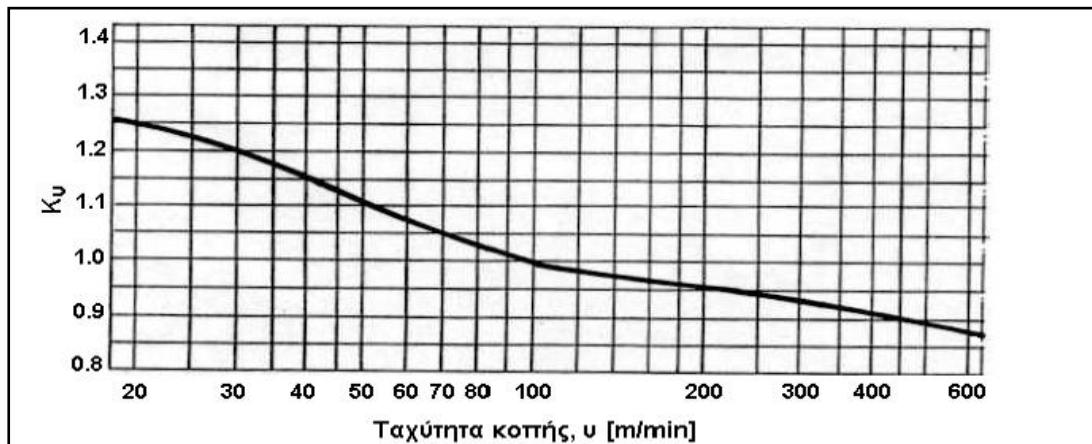
$$K_\gamma = 1 - \frac{\gamma - \gamma_o}{66.7}$$

όπου:

$\gamma$ : η τρέχουσα τιμή της γωνίας του χρησιμοποιούμενου κοπτικού εργαλείου.

$\gamma_o = +6^\circ$  για διαμήκη τόννευση χάλυβα ή  $+2^\circ$  για διαμήκη τόννευση χυτοσίδηρου.

- $K_v$ : συντελεστής διόρθωσης λόγω ταχύτητας κοπής στην περιοχή τιμών 20- 600 m/min, τιμές του συντελεστή λαμβάνονται από το Σχ. 9 συναρτήσει της πραγματικής ταχύτητας κοπής  $v$ .



Σχήμα 1.8 Υπολογισμός του συντελεστή διόρθωσης  $K_v$  [2].

- $K_\varepsilon$ : συντελεστής διόρθωσης λόγω υλικού κοπτικού εργαλείου, με τιμές 0.90-0.95 για κοπτικό εργαλείο από κεραμικό υλικό και 1 για κοπτικό εργαλείο από ταχυχάλυβα και σκληρομέταλλο.

- $K_\phi$ : συντελεστής διόρθωσης λόγω φθοράς του κοπτικού εργαλείου, με τιμές 1.30 - 1.50, ανάλογα με την έκταση του πεδίου φθοράς.
- ❖ Εμπειρικός υπολογισμός των υπολοίπων συνιστωσών δυνάμεων
  - Για τórνευση χαλύβων  
 $F_c : F_f : F_r = 10 : (4 \div 6) : (2 \div 3)$
  - Για τórνευση χυτοσιδήρων:  
 $F_c : F_f : F_r = 10 : (3 \div 6) : (1.5 \div 3)$

### 3. Μοντέλο υπολογισμού δυνάμεων για κοπές υψηλών ταχυτήτων

Πρόσφατα έχει δημοσιευθεί και μια τρίτη μέθοδος υπολογισμού της κύριας συνιστώσας της δύναμης κοπής. Σύμφωνα με αυτή την δημοσίευση η δύναμη κοπής  $F_c$  ισούται με την ειδική αντίσταση κοπής  $k_s$  πολλαπλασιασμένη με τον συντελεστή  $S$ :

$$F_c = k_s \cdot S$$

όπου  $S = a \cdot f$

Συνεπώς,  $F_c = k_s \cdot a \cdot f$

Η ειδική αντίσταση κοπής  $k_s$  εξαρτάται από το υλικό του κοπτικού εργαλείου, την πρόωση και την ταχύτητα κοπής. Μπορεί να υπολογιστεί, λοιπόν, μέσω της παρακάτω σχέσης:

$$k_s = k_1 \cdot f^{k_2} \cdot V_c^{k_3}$$

Κατά συνέπεια, η τελική σχέση μέσω της οποίας υπολογίζεται η δύναμη κοπής είναι η παρακάτω:

$$F_c = k_1 \cdot f^{k_2+1} \cdot V_c^{k_3} \cdot a$$

### 1.6 Ρυθμός αφαίρεσης υλικού

Ο ρυθμός αφαίρεσης υλικού (MRR - Material Removal Rate) είναι ο όγκος του αφαιρούμενου υλικού στην μονάδα του χρόνου ( $\text{mm}^3/\text{min}$ ). Παρατηρείται ότι για κάθε περιστροφή του τεμαχίου, αφαιρείται μια στρώση υλικού σχήματος δακτυλίου του οποίου το εμβαδό της επιφάνειας τομής αποτελείται από το γινόμενο της απόστασης που μετακινήθηκε το εργαλείο σε μια περιστροφή (πρόωση  $f$ ) και του βάθους κοπής  $a$ . Ο όγκος του δακτυλίου προκύπτει από το εμβαδόν της επιφάνειας τομής (δηλαδή  $f \cdot a$ ) της μέσης περιφέρειας του δακτυλίου (δηλαδή  $\pi D_{\text{avg}}$ , όπου  $D_{\text{avg}} = (D_a + D_f)/2$ ).

Για μικρά βάθη κοπής σε μεγάλες διαμέτρους τεμαχίων η μέση διάμετρος μπορεί να αντικατασταθεί από  $D_a$ . Η ταχύτητα περιστροφής του τεμαχίου είναι  $n$ . Συνεπώς, ο ρυθμός αφαίρεσης υλικού σε κάθε περιστροφή θα είναι:

$$\text{MRR} = (\pi)(D_{\text{avg}})(d)(f)(n) [\text{mm}^3/\text{min}]$$



Ο χρόνος κοπής  $t$  για ένα τεμάχιο μήκους  $l$  μπορεί να υπολογιστεί λαμβάνοντας υπόψη ότι το εργαλείο κινείται με ταχύτητα ίση με τον ρυθμό πρόωσης  $fn$ . Εφόσον η απόσταση που διανύεται είναι  $l$  mm, ο χρόνος κοπής είναι:

$$t = \frac{l}{fn}$$

Ο χρόνος αυτός δεν περιλαμβάνει τον χρόνο που απαιτείται για την προσέγγιση και την απομάκρυνση του εργαλείου από και προς το τεμάχιο. Λόγω του ότι ο νεκρός χρόνος είναι μη παραγωγικός και επιδρά έντονα στα συνολικά κόστος της κατεργασίας θεωρείται πολύ σημαντικός. Οι σημερινές εργαλειομηχανές είναι σχεδιασμένες ώστε να ελαχιστοποιούνται οι νεκροί χρόνοι [2,3].

### **1.7 Τα κοπτικά εργαλεία στην κατεργασία της τórνευσης**

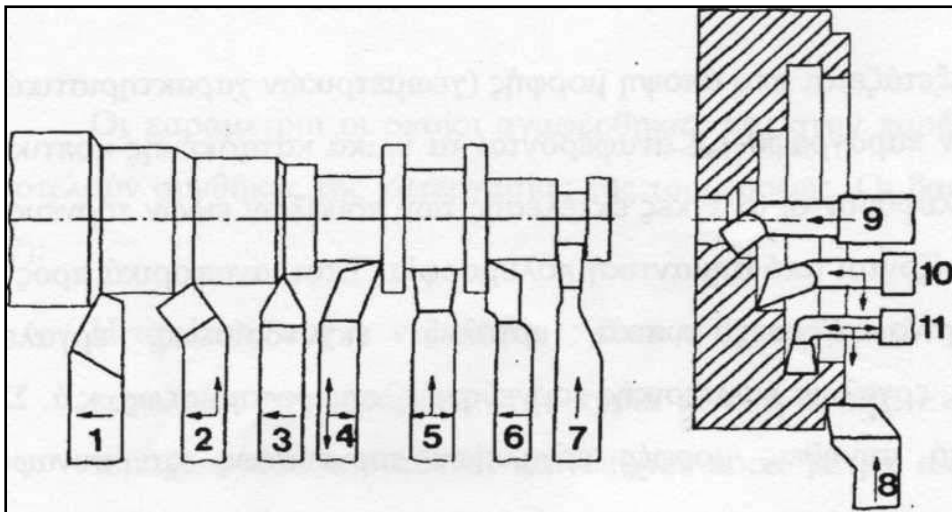
Κάθε κοπτικό εργαλείο, εξετάζεται από άποψη μορφής (γεωμετρικών χαρακτηριστικών) και υλικού κατασκευής του. Για να καλυφθούν οι ανάγκες εκτέλεσης των ποικίλων ειδών τórνευσης τα αντίστοιχα εργαλεία χαρακτηρίζονται από σημαντική πολυμορφία. Έτσι, αναφορικά προς τη φύση λειτουργίας τους διακρίνονται σε: τυπικά, εργαλεία εκχονδρίσεως, εργαλεία αποπερατώσεως, ειδικά εργαλεία, εργαλεία εσωτερικής τórνευσεως, σπειροτομήσεως, κ.ά. Στο σχήμα 1.9 δίδονται σχηματικά συνήθεις μορφές εργαλείων τórνευσεως και συναφείς εφαρμογές τους.

Ως επί το πλείστον, στην κατασκευή των εργαλείων τórνευσεως χρησιμοποιούνται οι ταχυχάλυβες και τα σκληρομέταλλα των εκάστοτε ενδεδειγμένων κοπτικών ιδιοτήτων και γεωμετρικών στοιχείων, χωρίς όμως δευτερευόντως να αποκλείονται σε ειδικές περιπτώσεις: οι ανθρακούχοι και κεκραμένοι χάλυβες εργαλείων, τα χυτευτά κράματα, τα κεραμικά πυριμάχων οξειδίων και το διαμάντι [2,3].

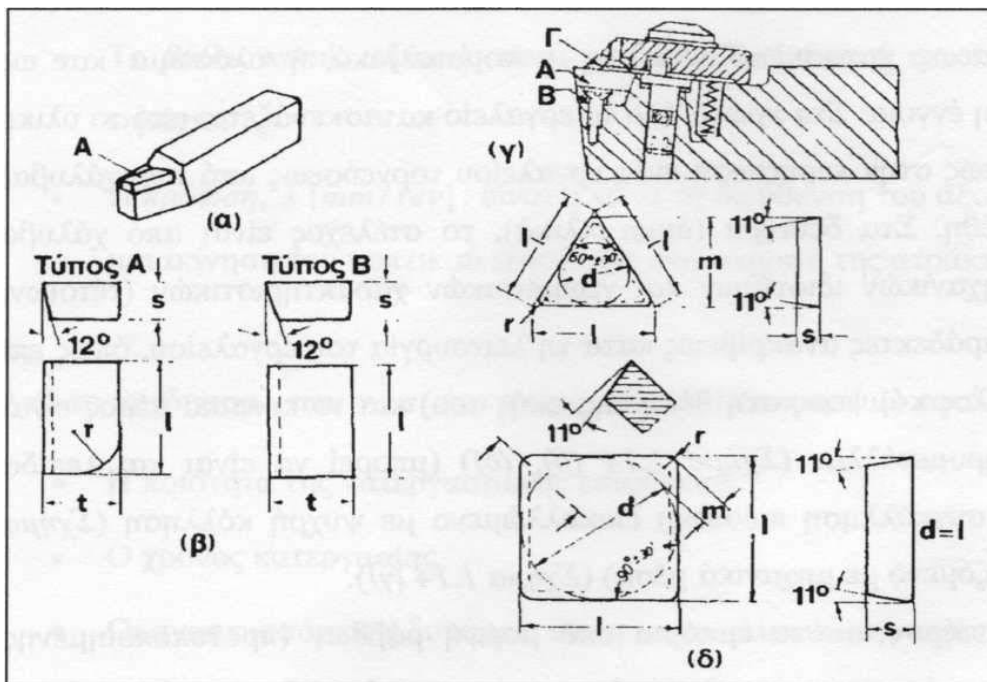
Τα εργαλεία τórνευσεως κατασκευάζονται ως μονομεταλλικά ή ολόσωμα και ως διμεταλλικά, με την ακόλουθη έννοια: στα πρώτα όλο το εργαλείο κατασκευάζεται από το υλικό του κοπτικού μέρους του (όπως στην περίπτωση ενός εργαλείου τórνευσεως από ταχυχάλυβα) και σε μικρά σχετικώς μεγέθη. Στα δεύτερα (διμεταλλικά), το στέλεχος είναι από χάλυβα κατασκευών κατάλληλων μηχανικών ιδιοτήτων και γεωμετρικών χαρακτηριστικών (τέτοιων, ώστε να μην προκύπτουν απαράδεκτες ανακρίβειες κατά τη λειτουργία του εργαλείου, όπως επί παραδείγματι απαράδεκτο βέλος κάμψεως στη θέση της ακής του) και το κοπτικό μέρος είναι κατά κανόνα πλακίδιο σκληρομετάλλου (σχήμα 1.10 (β), (δ)) (μπορεί να είναι και λεπίδα ταχυχάλυβα στερεωμένη με συγκόλληση πίεσεως) επικολώμενο με ψυχρή κόλληση (σχήμα 1.10 (α)) ή ένθετο (προσαρμοζόμενο με μηχανικά μέσα) (σχήμα 1.10 (γ)).

Οι ταχυχάλυβες προσφέρονται στο εμπόριο υπό μορφή ράβδων (προτυποποιημένης μορφής και διαστάσεων διατομής, που μπορεί να είναι τετραγωνική, ορθογωνική, κυκλική, τριγωνική ή τραπεζοειδής) έτοιμων προς χρήση (έχουν, δηλαδή, υποστεί την ενδεδειγμένη βαφή και διπλή επαναφορά). Μετά από προσεκτικό τρόχισμα σε ειδικό τροχιστικό μηχάνημα στην επιθυμητή μορφή, στερεώνονται στα στελέχη (μανέλλες) [2,3].

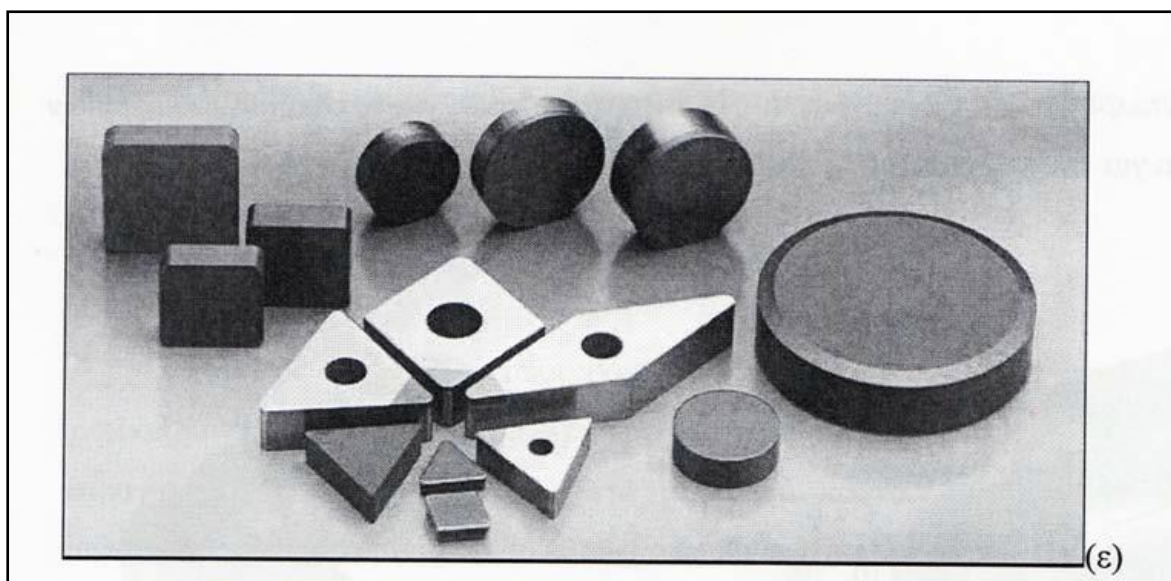
Τα σκληρομέταλλα προσφέρονται στο εμπόριο με τη μορφή πλακιδίων σε προτυποποιημένα σχήματα και μεγέθη (σχήμα 1.11).



**Σχήμα 1.9** Σχηματική παράσταση τυποποιημένων μορφών εργαλείων τριβής και παράδειγμα εφαρμογής τους: 1. ευθύ εργαλείο εκχονδρίσεως, 2. κεκαμμένο εκχονδρίσεως, 3. οξύ αποπερατώσεως 4. κεκαμμένο ειδικής μορφής, 5. πλατύ (ορθογωνικό) αποπερατώσεως, 6. ξεθυμάσματος (γωνιάσματος), 7. αποκοπής, 8. μετωπικής τριβής (προσώπου), 9. εσωτερικής τριβής (διαμπερή κοίλα), 10. τριβής τυφλών κοίλων, 11. ορθογωνικό εσωτερικής αυλακώσεως [2].



**Σχήμα 1.10** Μορφές προτυποποιημένων πλακιδίων από σκληρομέταλλο και τρόποι συγκρατήσεως του στο στέλεχος του κοπτικού εργαλείου. α) Εργαλείο με επικολλώμενο πλακίδιο Α. β) Μορφές επικολλώμενων πλακιδίων, γ) Συγκράτηση ένθετου πλακιδίου: Α. πλακίδιο, Β. Προσθήκη, Γ. γρεζοθραύστης, δ) μορφές ένθετων πλακιδίων [2].



**Σχήμα 1.11** Διάφορες μορφές πλακιδίων σκληρομετάλλων [2].

Ύστερα από πολλές οικονομοτεχνικές μελέτες που έχουν διεξαχθεί μέχρι σήμερα, αλλά και από τη συσσωρευμένη εμπειρία, έχει πλέον καθιερωθεί το γεγονός ότι τα εργαλεία με ένθετα πλακίδια πλεονεκτούν σημαντικά έναντι των εργαλείων με επικολλώμενα πλακίδια. Και τούτο, γιατί η χαλάρωση του πλακιδίου και απλή περιστροφή του αρκούν για την αλλαγή της κόψης του που τυχόν έχει φθαρεί (έτσι μπορούν διαδοχικά να χρησιμοποιηθούν και οι έξι ή και οι οκτώ κόψεις του πλακιδίου, αν το πλακίδιο είναι αντίστοιχα τριγωνικό ή τετραγωνικό και έχει μηδενική γωνία ελευθερίας, αλλιώς οι τρεις ή οι τέσσερις κόψεις του), χωρίς να απαιτείται αφαίρεση του εργαλείου από την εργαλειομηχανή και αποκόλληση, τρύχιση και επικόλληση του πλακιδίου, όπως συμβαίνει στα εργαλεία με επικολλώμενα πλακίδια. Επιπλέον, τα ένθετα πλακίδια δεν διατρέχουν τον κίνδυνο καταστροφής, ο οποίος επαπειλεί τα επικολλώμενα κατά την κόλληση τους στο στέλεχος [3].

Τέλος, να αναφερθούμε σε μια ειδική κατηγορία κοπτικών εργαλείων, τους κυβικούς βοριοιτρίτες (CBN). Ο κυβικός βοριοιτρίτης παράγεται με πυροσυσσώματωση υπό υψηλή θερμοκρασία και υψηλή πίεση σωματιδίων κυβικού βοριοιτρίτη και ενός συνδετικού υλικού. Το αρχικό υλικό είναι κόκκοι που χρησιμοποιούνται για την λείανση κραμάτων φερριτικών μετάλλων. Τα πλακίδια από κυβικό βοριοιτρίτη επικοκολλούνται πάνω σε καρβίδια για να μπορέσουν να κρατηθούν από τις μανέλλες [17].

Στη σκληρότητα έρχονται δεύτερα μετά τα εργαλεία από διαμάντι. Οι μηχανικές και θερμικές τους ιδιότητες το κάνουν ένα υπέρ-εργαλείο για δυσκατέργαστα, σκληρά φερριτικά κράματα. Η σκληρότητά τους στους 704°C είναι ακόμα καλύτερη από αυτή των καρβιδίων και των κεραμικών σε θερμοκρασία περιβάλλοντος. Σε υψηλή θερμοκρασία, τα εργαλεία από καρβίδιο τείνουν να μαλακώσουν με ρυθμό ίδιο με αυτό του προς κατεργασία υλικού. Η συμπεριφορά αυτή σε συνδυασμό με τη χημική αντίδραση (οξειδωση) περιορίζουν τη χρήση τους σε θερμοκρασίες κάτω από 704°C.

Η θερμική ανθεκτικότητα και αγωγιμότητα των εργαλείων κυβικού βοριοιτρίτη τα καθιστά κατάλληλα για χρήση σε υψηλές προώσεις και ταχύτητες κοπής, όπου τα απλά εργαλεία δεν μπορούν να αντεπεξέλθουν. Σε γενικές γραμμές, τα συγκεκριμένα εργαλεία χρησιμοποιούνται για την λείανση εξαρτημάτων, όπως γρανάζια από επιβελτιωμένους χάλυβες ή στην κατεργασία υπέρσκληρων κραμάτων όπως τιτανίου, νικελίου και κεραμικών υλικών [17].

Γενικά, για τις κατεργασίες των κεραμικών με βάση μελέτες δεν χρησιμοποιείται ψυκτικό υγρό είτε για οικολογικούς λόγους (ξηρά κοπή) είτε για λόγους ευκολίας στην κατεργασία. Για τα κεραμικά υλικά, για παράδειγμα, πρέπει κατά την κατεργασία τους να μην χρησιμοποιείται ψυκτικό υγρό, πρώτον για να μπορούν μέσω των υψηλών θερμοκρασιών που θα αναπτυχθούν στην διεπιφάνεια εργαλείου/τεμαχίου να γίνουν ευκατέργαστα και δεύτερον, για να μην υποστούν κάποιο θερμικό αιφνιδιασμό, με αποτέλεσμα τη δημιουργία ρωγμών ή την ολοκληρωτική τους θραύση. Επομένως, τα μοναδικά εργαλεία που μπορούν να χρησιμοποιηθούν για τις παραπάνω περιπτώσεις είναι τα εργαλεία κυβικού βοριονιτρίτη [17].

### **1.8 Τα υγρά κοπής στην κατεργασία της κοπής**

Τα υγρά κοπής χρησιμοποιούνται συχνά στις διάφορες κατεργασίες, κατάλληλα προσαγόμενα στην περιοχή κοπής την οποία διαβρέχουν, σε αντίθεση με τη λεγόμενη ξηρή κοπή που εκτελείται χωρίς υγρό κοπής. Η δράση του υγρού κοπής οφείλεται, όπως αναφέρεται παρακάτω, σε ορισμένες χαρακτηριστικές του ιδιότητες και εκδηλώνεται διαφορετικά και ανάλογα με το υλικό του κατεργαζόμενου τεμαχίου, το υλικό και τη γεωμετρία του κοπτικού εργαλείου και με τις συνθήκες κοπής [3].

Έτσι, ένα υγρό κοπής μπορεί να επενεργήσει είτε ως λιπαντικό είτε ως ψυκτικό και συχνά ως λιπαντικό και ως ψυκτικό μαζί. Η γνώση των δράσεων αυτών του υγρού κοπής μας είναι πολύ χρήσιμη, γιατί μας διευκολύνει στην επιτυχή εκλογή του στις διάφορες εφαρμογές. Πιο αναλυτικά, με τη χρησιμοποίηση των υγρών κοπής μπορούμε να επιτύχουμε:

- Μείωση στο μέσο φαινόμενο συντελεστή τριβής στη διεπιφάνεια αποβλίττου-εργαλείου, αλλά και στο μέσο συντελεστή τριβής στην επιφάνεια επαφής εργαλείου και τεμαχίου (σε περιπτώσεις, που εμφανίζεται σημαντική σχετικά ζώνη φθοράς) ως απόρροια της λιπαντικής δράσεως του υγρού κοπής.
- Ελάττωση της θερμοκρασίας του εργαλείου, του τεμαχίου και του αποβλίττου μέσω απαγωγής μέρους της θερμότητας (που παράγεται κατά την κοπή) ως αποτέλεσμα της ψυκτικής δράσεως του υγρού κοπής.
- Μείωση της φθοράς του κοπτικού εργαλείου (και αύξηση της ζωής του) έμμεσα, ως αποτέλεσμα της ελαττώσεως του συντελεστή τριβής και της ψύξεως
- Βελτίωση της τραχύτητας της κατεργασμένης επιφάνειας του τεμαχίου.
- Μείωση στις δυνάμεις και στην ισχύ κοπής.
- Περιορισμό στις στρεβλώσεις του τεμαχίου, οι οποίες προκαλούνται λόγω υψηλών τοπικών διαφορών της θερμοκρασίας, που αναπτύσσεται.
- Ευκολία στο χειρισμό των έτοιμων τεμαχίων. Με τη χρήση κατάλληλου ψυκτικού κοπής είναι δυνατή η διατήρηση ομοιόμορφης, σε κάποιο βαθμό, θερμοκρασίας του τεμαχίου σε χαμηλότερη στάθμη.
- Προστασία του τεμαχίου και των μερών της εργαλειομηχανής, που έρχονται σε επαφή υγρό κοπής, από οξείδωση ή διάβρωση.
- Απομάκρυνση μικρών τεμαχίων αποβλίττου και άλλων σωμάτων (μικροτμηματιδίων, που προκύπτουν από τη φθορά του εργαλείου, όπως για παράδειγμα από φθορά λόγω αποξέσεως). Αυτό προφυλάσσει από τραυματισμούς την κατεργασμένη επιφάνεια του τεμαχίου και συνεπώς συμβάλλει στη διατήρηση της επιθυμητής τραχύτητας της.

Το υγρό κοπής πρέπει να έχει ικανοποιητικές ιδιότητες διαβρεκτικότητας επιφανειών, με τις οποίες αυτό έρχεται σε επαφή. Επιπλέον οφείλει να είναι (και να παραμένει), φυσικώς και χημικώς ευσταθές και να διατηρεί την αποτελεσματικότητά του κατά όλη την ωφέλιμη ζωή του, που μπορεί να φθάσει και τους 12 μήνες. Τέλος, πρέπει να είναι οικονομικό. Δηλαδή, το κόστος κτήσεως του υγρού κοπής, μαζί με εκείνο της εγκαταστάσεως και λειτουργίας του συστήματος παροχής του, οφείλει να αντισταθμίζει τουλάχιστον στις ωφέλειες, που προκύπτουν από τη χρήση του στις διάφορες κατεργασίες.

Οι βασικές δράσεις του υγρού κοπής είναι οι πρώτες δύο εκ των αναφερόμενων, όμως, άλλες από τις επενέργειες του μπορούν να θεωρηθούν ως επακόλουθο των πρώτων.

Τα υγρά κοπής χρησιμοποιούμενα στο μηχανουργείο εμφανίζουν εν γένει και ορισμένες ανεπιθύμητες παρενέργειες, που είναι δυνατό ή να καταστήσουν αδύνατη τη χρησιμοποίησή τους ή να περιορίσουν το πεδίο εφαρμογών τους [3]. Ως τέτοιες βασικές παρενέργειες παραθέτονται οι ακόλουθες:

A. Παρενέργειες στην υγεία του τεχνίτη (για παράδειγμα αναπνευστικά προβλήματα, δερματίτιδες, καρκίνοι) οφειλόμενες σε τοξικούς ατμούς, δυσάρεστες οσμές, αναθυμιάσεις και άμεση επαφή των υγρών κοπής με τον εργάτη.

B. Επιδράσεις στο υλικό του τεμαχίου ή σε υλικά της εργαλειομηχανής, που εκδηλώνονται συνήθως με οξείδωση ή διάβρωση. Υγρά κοπής για παράδειγμα με πρόσθετα υψηλής πίεσεως, τα οποία περιέχουν θείο, προσβάλλουν τα κράματα του χαλκού, που τυχόν κατεργάζονται, όπως και τα έδρανα της εργαλειομηχανής, τα οποία έχουν ως βάση το χαλκό και είναι προσιτά στο υγρό κοπής.

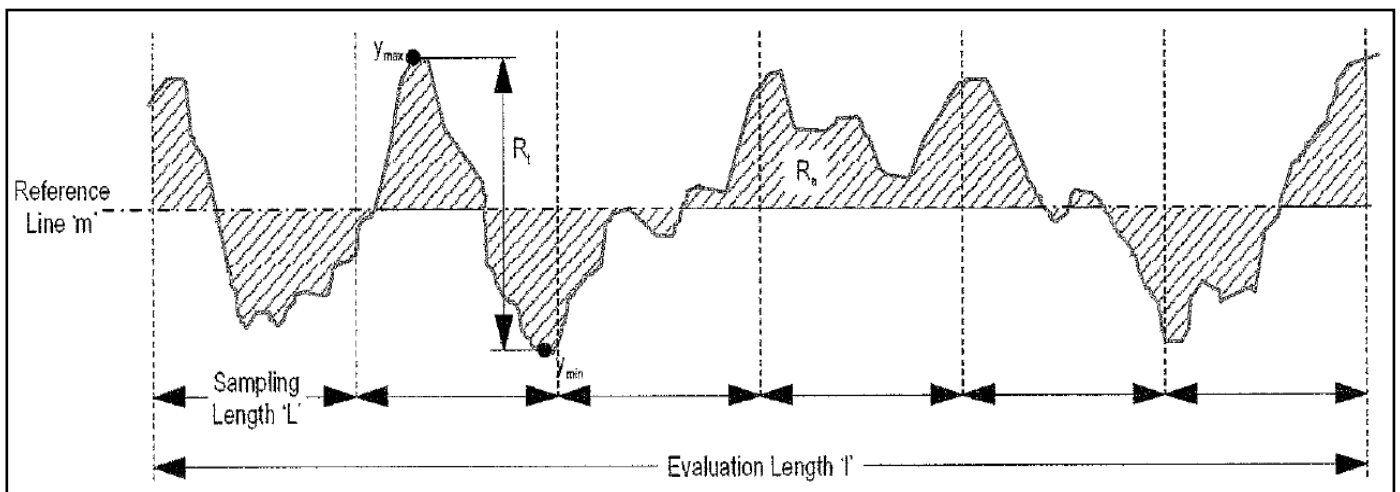
Οι ανεπιθύμητες παρενέργειες του υγρού κοπής μπορούν να περιοριστούν ή και να εξαφανιστούν πλήρως με την εκλογή κάθε φορά υγρού κοπής με τις ενδεδειγμένες ιδιότητες (όσον το δυνατόν φιλικότερο προς το περιβάλλον και τον χρήστη) ή με την υιοθέτηση σύγχρονων τεχνολογιών και κατάλληλων συνθηκών κατεργασιών κοπής (υπέρσκληρα υλικά κοπτικών εργαλείων, εργαλειομηχανές) όπου εφαρμόζεται η ξηρή κοπή [3].

## 1.9 Επιφανειακή τραχύτητα έπειτα από τόννευση

Τα όρια των στερεών σωμάτων είναι σχηματισμένα από τις επιφάνειές τους. Στη μηχανική αυτά τα όρια περιγράφονται γεωμετρικά σε σχεδιαστικά μοντέλα ως τέλειες επιφάνειες (συχνά αναφέρονται ως γεωμετρικές επιφάνειες). Τα παραγόμενα εξαρτήματα όμως εμφανίζουν πάντα απόκλιση από τις αντίστοιχες γεωμετρικές επιφάνειές τους. Η επιφανειακή τραχύτητα είναι μια όψη των αποκλίσεων αυτών. Το μέγεθος, η μορφή και ο κυματισμός είναι άλλες όψεις αυτής της απόκλισης.

Για να μπορέσουμε να μετρήσουμε την επιφανειακή τραχύτητα, ορίζουμε ένα προφίλ μετρώντας τις αποκλίσεις που είναι κάθετες στην γεωμετρική επιφάνεια. Η γραμμή  $m$  λοιπόν χρησιμοποιείται ως σημείο αναφοράς για τον υπολογισμό κάποιων παραμέτρων της τοπογραφίας της επιφάνειας. Είναι υπολογισμένη ως μια προσέγγιση με χρήση της μεθόδου ελαχίστων τετραγώνων των σημείων που είναι αποτυπωμένα κατά μήκος του μήκους μέτρησης. Επομένως η συνολική περιοχή άνω της γραμμής αναφοράς που περικλύεται από τη καταγεγραμμένη τροχιά είναι ίση με την αντίστοιχη από κάτω [4].

Οι κύριες παράμετροι της τοπογραφίας της επιφάνειας είναι το  $R_t$  και το  $R_a$  (βλ.σχήμα 1.12) όπου:



Σχήμα 1.12 Επιφανειακή Τραχύτητα [4].

Η παράμετρος  $R_t$  μετράει την κάθετη απόσταση μεταξύ του υψηλότερου ( $y_{max}$ ) και του χαμηλότερου ( $y_{min}$ ) σημείου κατά το μήκος μέτρησης.

Οπότε:  $R_t = y_{max} - y_{min}$

Η παράμετρος  $R_a$  μετράει τη μέση απόκλιση (κατά απόλυτη τιμή) των μετρούμενων τιμών από τη γραμμή αναφοράς.

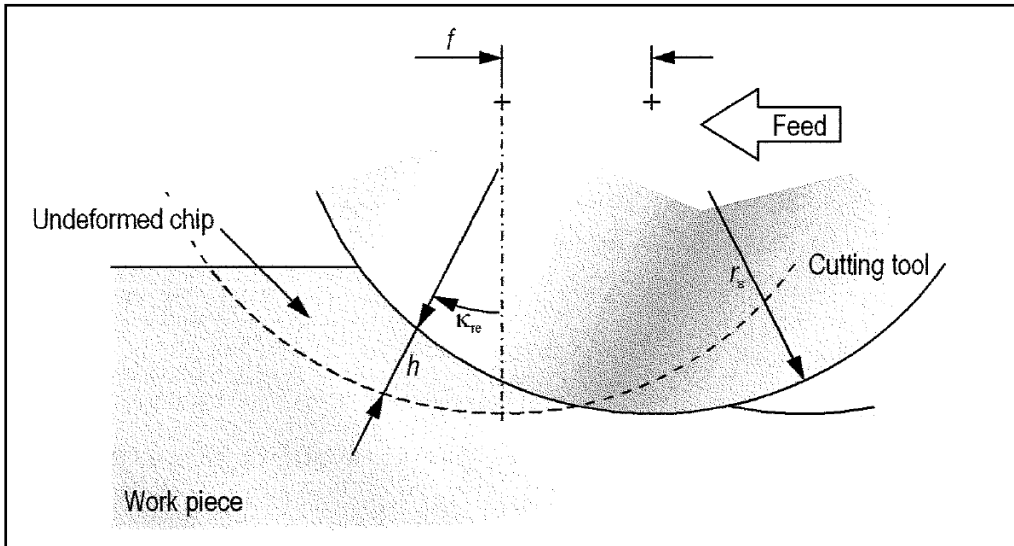
$$\text{Οπότε: } R_a = \frac{1}{l \cdot \int (|y|) dx}$$

Προσεγγιστικά μπορούμε να πούμε για πεπερασμένα σημεία ότι:

$$R_a = \frac{|y_1| \cdot |y_2| \cdot |y_3| \cdots |y_n|}{n}$$

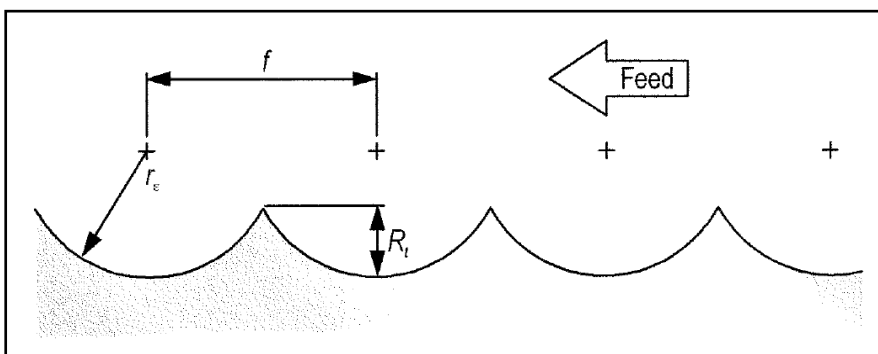
### ΔΗΜΙΟΥΡΓΙΑ ΚΑΤΕΡΓΑΣΜΕΝΗΣ ΕΠΙΦΑΝΕΙΑΣ

Κατά τη τόννευση το κατεργαζόμενο κομμάτι περιστρέφεται υπό το κινούμενο κοπτικό εργαλείο, γεγονός που οδηγεί το κοπτικό εργαλείο να ακολουθήσει μια ελλειπτική τροχιά πάνω στη κατεργαζόμενη επιφάνεια. Με κάθε περιστροφή του τεμαχίου δημιουργείται στην κατεργαζόμενη επιφάνεια ένα αποτύπωμα του προφίλ του κοπτικού εργαλείου. Κάθε αποτύπωμα απέχει από το επόμενο αποτύπωμα κατά απόσταση ίση με την πρόωση ( $f$ ) [4].



**Σχήμα 1.13** Διαστάσεις και μορφή αποβλήτου πριν παραμορφωθεί από την κοπή του (απαραμόρφωτο απόβλητο) [4].

Με γνωστό το προφίλ ενός κοπτικού εργαλείου (συνήθως ακτινικό) είναι δυνατός ο υπολογισμός του προφίλ της κατεργασμένης επιφάνειας με τη βοήθεια της συχνά χρησιμοποιούμενης παραμέτρου  $R_t$  (peak-to-valley) όπως φαίνεται παρακάτω:



**Σχήμα 1.14** Δημιουργία επιφανειακής τραχύτητας κατά την τόννευση [4].

Οι μαθηματικοί υπολογισμοί προϋποθέτουν το προφίλ του κοπτικού εργαλείου να είναι απόλυτα ακτινικό.

### **1.10 Τόρνευση σκληρών υλικών (hard turning)**

Ως hard turning ορίζεται η κατεργασία σημειακής κοπής κομματιών που έχουν τιμές σκληρότητας μεγαλύτερες των 45 HRC. Συνήθως, ωστόσο, κομμάτια που υπόκεινται hard turning έχουν σκληρότητα που κυμαίνεται στο πεδίο 58-68 HRC. Η κατεργασία του hard turning είναι αρκετά όμοια με της συμβατικής τόρνευσης οπότε η πραγματοποίησή της σε ένα φυσιολογικό εργοστασιακό περιβάλλον είναι δυνατή με μικρές λειτουργικές αλλαγές (εφόσον υπάρχουν τα σωστά εργαλεία).

Παρά το γεγονός ότι το στήσιμο του hard turning είναι παρόμοιο με αυτό της κανονικής (συμβατικής) τόρνευσης (soft turning) πρέπει να σημειωθεί ότι οι συνθήκες κοπής είναι σαφώς διαφορετικές. Μιλώντας με παραμέτρους κατεργασίας, η κατεργασία hard turning χαρακτηρίζεται από χαμηλή πρόωση ( $f$ ) και μικρό βάθος κοπής ( $a_p$ ).

Επίσης τυπικό χαρακτηριστικό για τη κατεργασία hard turning (επειδή αποτελεί συνήθως κατεργασία ξηράς κοπής) είναι η υψηλή έκλυση θερμότητας. Η μέγιστη θερμοκρασία που αναπτύσσεται κατά το hard turning βρίσκεται στο σημείο επαφής αποβλίττου και εργαλείου κοπής και μπορεί να φθάσει σε θερμοκρασίες μεγαλύτερες των 750 βαθμών κελσίου. Υπάρχουν 2 βασικές πηγές αυτής της θερμότητας: Η παραμόρφωση του αποβλίττου στην πρωτεύουσα ζώνη διάτμησης και η τριβή στην διεπαφή αποβλίττου και κοπτικού εργαλείου. Υψηλότερες ταχύτητες κοπής προκαλούν μεγαλύτερη εισαγωγή ενέργειας ανά μονάδα χρόνου οπότε αυξάνουν τις θερμοκρασίες κατά τη κοπή, ενώ παράλληλα, η υψηλή θερμική αγωγιμότητα ενός κοπτικού εργαλείου επιτρέπει στην απορρόφηση περισσότερης ενέργειας από την επιφάνεια επαφής αποβλίττου - κοπτικού εργαλείου οπότε οι θερμοκρασίες μειώνονται [4].

Η κατεργασία hard turning πραγματοποιείται με καλύτερο τρόπο όταν τα κοπτικά εργαλεία είναι κατασκευασμένα είτε από CBN (Cubic Boron Nitride) είτε από Cermet (CERamic και METallic) ή από κεραμικά υλικά. Επειδή το hard turning είναι σημειακής κοπής, ένα σημαντικό κέρδος από την κατεργασία αυτή είναι η δυνατότητα να δημιουργήσουμε περιγράμματα (contours) και να παράγουμε πολύπλοκα σχήματα εκμεταλλευόμενοι την ικανότητα κίνησης των καινούργιων κοπτικών εργαλείων. Για πολλές εφαρμογές, η κοπή με κοπτικό CBN είναι η καλύτερη επιλογή. Ωστόσο, για κάποιες περιπτώσεις κατεργασίας επιλέγονται και τα κοπτικά από Cermet ή Κεραμικά υλικά.

Η τόρνευση κομματιών υψηλής σκληρότητας μπορεί να χρησιμοποιηθεί ως μια κατεργασία που αντικαθιστά τη λείανση, ή μπορεί επίσης να χρησιμοποιηθεί ως προ-λείαντική προετοιμασία. Ωστόσο, σπάνια χρησιμοποιείται για κατασκευή εξαρτημάτων που απαιτούν μεγάλη ακρίβεια. Γενικά χρησιμοποιείται για εξαρτήματα όπου οι απαιτήσεις ακρίβειας και φινιρίσματος της επιφάνειας δεν είναι υψηλές. Σε υψηλής ακρίβειας εφαρμογές η λείανση παραμένει ακόμη η συχνότερα χρησιμοποιούμενη κατεργασία καθώς η τελική επιφάνεια που προκύπτει από την κατεργασία hard turning δεν μπορεί ακόμη να ανταγωνισθεί αυτήν των μηχανών λείανσης υψηλής ακρίβειας. Επίσης, η λείανση έχει - και πάντα θα έχει - θέση στις κατεργασίες, καθώς δεν είναι όλα τα εξαρτήματα σε θέση να υποστούν hard turning λόγω απαιτήσεων στις ανοχές και στην πιστότητα επιφανείας του εκάστοτε εξαρτήματος [4].

Σε καθημερινή βάση κομμάτια υπόκεινται στη κατεργασία hard turning στους παρακάτω τομείς της βιομηχανίας:

- Στην αυτοκινητοβιομηχανία
- Στις κατασκευές ρουλεμάν
- Στην ναυτιλία
- Στην κατασκευή εμβόλων και μητρών (punch and die)
- Στην κατασκευή καλουπιών



- Στην κατασκευή εργαλειομηχανών
- Στην αεροναυπηγική

ενώ κάθε μέρα προστίθενται και άλλοι τομείς [4].

Τα συνήθη υλικά, τα οποία χρησιμοποιούνται καθημερινά, περιλαμβάνονται στις παρακάτω κατηγορίες:

- Κράμματα χάλυβα
  - Χάλυβες ρουλεμάν
  - Χαλύβδινα εργαλεία
  - Εργαλεία χάραξης/κοπής
  - Χάλυβες καλουπιών
  - Χάλυβες με επιφανειακή σκλήρυνση
- και άλλα πολλά [4].

Κάποιοι από τους λόγους που μπορεί να προτιμηθεί η χρήση της σκληρής τόνρευσης έναντι της λείανσης είναι ότι:

- Μπορεί να πραγματοποιηθεί συμβατική τόνρευση ή hard turning στην ίδια μηχανή
  - Χρειάζεται λιγότερος χώρος συνολικά
  - Είναι μια επένδυση με χαμηλότερο συνολικό κόστος καθώς οι τόρνοι είναι σχετικά λιγότερο ακριβοί να αγοραστούν. Οι υψηλότερων απαιτήσεων λειαντικές μηχανές μπορούν να έχουν πολλαπλές χρήσεις, αλλά συχνά είναι πιο ακριβοί στην χρήση τους λόγω του στησίματος και των χρόνων κύκλου και απαιτούν περισσότερο εξοπλισμό υποστήριξης
  - Ο ρυθμός αφαίρεσης μετάλλου είναι 4-6 φορές μεγαλύτερος από αυτόν της λείανσης
  - Μπορεί να πραγματοποιηθεί τόνρευση σύνθετων περιγραμμάτων (contours). Οι τόρνοι μπορούν να κατεργασθούν μικρό αριθμό εξαρτημάτων και πολύπλοκες μορφές πολύ αποτελεσματικότερα από τις μηχανές λείανσης. Η σημειακή τόνρευση πολύπλοκων contours μπορεί να επιτευχθεί, εξαλείφοντας τους ακριβούς τροχούς μορφής.
  - Μπορούν να πραγματοποιηθούν πολλαπλές κατεργασίες με ένα μόνο στησιμο
  - Επιτρέπει τη δημιουργία τραχύτητας επιφανείαςτης τάξης του μικρομέτρου
  - Όσον αφορά σε περιβαλλοντικά θέματα, οι τόρνοι παράγουν απόβλιττο, το οποίο είναι λιγότερο ακριβό να διαχειριστούμε από τη μορφή του αποβλίττου από τις λειαντικές μηχανές (swarf)
- [4]

## ΚΕΦΑΛΑΙΟ 2<sup>ο</sup> ΤΑ ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

### 2.1 Η ιστορία των τεχνητών νευρωνικών δικτύων

Οι πρώτες σημαντικές εργασίες στα Τεχνητά Νευρωνικά Δίκτυα (ΤΝΔ) ξεκίνησαν να παρατίθενται κατά το 1940. Πιο συγκεκριμένα το 1943, οι McCulloch και Pitts παρέθεσαν τη μαθηματική αναπαράσταση των νευρωνικών διεργασιών έπειτα από μελέτη ενός μοντέλου του βασικού κυττάρου του ανθρώπινου εγκεφάλου. Κατά το 1957 εκδόθηκε μια τεχνική αναφορά με την περιγραφή του απλοποιημένου νευρώνα perceptron του Rosenblatt. Παράλληλα πραγματοποιούνταν η έρευνα του Widrow πάνω στον αλγόριθμο ελαχίστων τετραγώνων για το μοντέλο Adaline/Madaline το οποίο περιγραφόταν από μια μονάδα ή ένα μοναδικό στρώμα. Στα χρόνια που ακολούθησαν, η γνώση πάνω στο συγκεκριμένο αντικείμενο κωδικοποιήθηκε σε ένα ξεχωριστό γνωστικό πεδίο, αυτό των Τεχνητών Νευρωνικών Δικτύων [5,6].

Το μεγαλύτερο πλήγμα στην πορεία των νευρωνικών δικτύων αποτέλεσε η δημοσίευση του βιβλίου "Perceptrons" των Minsky και Papert το 1969 όπου παρουσιάζονταν τα αρνητικά σημεία του απλοποιημένου μοντέλου νευρώνα (perceptron) με αποτέλεσμα οι χρηματοδοτήσεις στις έρευνες της εποχής να κατευθυνθούν κατά το μεγαλύτερο μέρος τους σε κλάδους άλλων γνωστικών αντικειμένων. Το ενδιαφέρον για τα τεχνητά νευρωνικά δίκτυα ελαττώθηκε αλλά δεν εξαλείφθηκε.

Στις επόμενες δεκαετίες τα μειονεκτήματα του perceptron ξεπεράστηκαν με την επινόηση των πολυστρωματικών δικτύων πρόσθιας τροφοδότησης (multi layer perceptron - MLP) και του αντίστοιχου αλγορίθμου εκπαίδευσης οπίσθιας τροφοδότησης (1985) ενώ ταυτόχρονα αναπτύχθηκαν τα δίκτυα αυτοοργάνωσης (self organizing maps - SOM) και η ανταγωνιστική μάθηση (1976), τα δίκτυα ακτινικών συναρτήσεων βάσης (radial basis function - RBF) και το δίκτυο Hopfield (1982). Οι παραπάνω εξελίξεις ήταν και το αποτέλεσμα της ανανέωσης του γενικότερου ενδιαφέροντος για τα νευρωνικά δίκτυα και ιδιαίτερα τα πολυστρωματικά τα οποία πλέον μπορούσαν να αντιμετωπίσουν προβλήματα που το perceptron αδυνατούσε να επιλύσει. Αναπτύχθηκαν νέες αρχιτεκτονικές διασύνδεσης όπως τα επαναληπτικά δίκτυα ενώ δημιουργήθηκε και η ιδέα της μάθησης χωρίς επίβλεψη [5,6].

Σήμερα τα ΤΝΔ έχουν αναπτυχθεί σε μεγάλο βαθμό χάρη στη συστηματική εκτίμηση των δυνατοτήτων και των περιορισμών τους και συνεχίζουν να αναπτύσσονται σε συνδυασμό με άλλες τεχνολογίες όπως οι Γενετικοί Αλγόριθμοι, τα Ασαφή και τα Έμπειρα Συστήματα. Ένας από τους κυριότερους παράγοντες στην εξέλιξη των ΤΝΔ είναι η αυξανόμενη δύναμη των υπολογιστικών πόρων, κυρίως ισχύς επεξεργαστών, μνήμη, παράλληλη επεξεργασία κλπ. [5,6]

### 2.2 Τι είναι τα τεχνητά νευρωνικά δίκτυα

Τα τεχνητά νευρωνικά δίκτυα είναι συστήματα τα οποία περιέχουν έναν μεγάλο αριθμό επεξεργαστών που ονομάζονται **νευρώνες**. Με τον όρο "τεχνητά" εννοούμε ότι οι υπολογισμοί δεν πραγματοποιούνται από βιολογικά στοιχεία αλλά με την χρήση μοντέλων εμπνευσμένων από τη βιολογία. Τα ΤΝΔ είναι ικανά να προσομοιώνουν κάποιες τουλάχιστον από τις λειτουργίες των πραγματικών νευρωνικών δικτύων όπως π.χ. να "μαθαίνουν" μέσα από παραδείγματα και να ανταποκρίνονται σε αντίστοιχα **ερεθίσματα** [5,6].

Όπως αναφέραμε, η **αρχιτεκτονική** τους δομή αποτελείται από ένα πλήθος διασυνδεδεμένων μονάδων - **τεχνητοί νευρώνες**. Κάθε μονάδα χαρακτηρίζεται από **εισόδους** και **εξόδους** και υλοποιεί τοπικά έναν απλό υπολογισμό. Κάθε σύνδεση μεταξύ δύο μονάδων χαρακτηρίζεται από μια **τιμή βάρους**. Οι τιμές των βαρών

των συνδέσεων αποτελούν τη **γνώση** που είναι αποθηκευμένη στο δίκτυο και καθορίζουν τη λειτουργικότητά του. Η έξοδος κάθε μονάδας καθορίζεται από τον τύπο της μονάδας, τη διασύνδεση με τις υπόλοιπες μονάδες και πιθανώς κάποιες εξωτερικές εισόδους. Συνήθως, ένα δίκτυο αναπτύσσει μια συνολική λειτουργικότητα μέσω μιας μορφής **εκπαίδευσης** όπου οι τιμές των βαρών παίρνουν τις βέλτιστες για τη λειτουργία αυτή τιμές. Μπορούμε να πούμε ότι η συνολική λειτουργικότητα ενός ΤΝΔ καθορίζεται από την τοπολογία του δικτύου, τα χαρακτηριστικά των νευρώνων, τη μέθοδο εκπαίδευσης και από τα δεδομένα με τα οποία γίνεται αυτή η εκπαίδευση [5,6].

### **2.3 Δυνατότητες και πλεονεκτήματα των τεχνητών νευρωνικών δικτύων**

Τα νευρωνικά δίκτυα έχουν την ικανότητα να εξάγουν πρότυπα και να παρατηρήσουν τάσεις που είναι αρκετά περίπλοκες για να παρατηρηθούν είτε από τον άνθρωπο είτε από άλλα υπολογιστικά συστήματα [7]. Κάποιες από τις χρήσιμες ιδιότητες των ΤΝΔ είναι οι εξής:

#### **Μη γραμμικότητα**

Οι τεχνητοί νευρώνες μπορούν να είναι είτε γραμμικοί είτε μη γραμμικοί, πράγμα που συνεπάγεται ότι ένα νευρωνικό δίκτυο αποτελούμενο από μη γραμμικούς νευρώνες είναι και αυτό από τη φύση του μη γραμμικό. Η μη γραμμικότητα αυτή είναι κατανεμημένη σε όλη την έκταση του δικτύου και είναι απαραίτητη για την εκπαίδευση του δικτύου σε φυσικούς μηχανισμούς που οι εισοδοί των οποίων δεν είναι γραμμικές (π.χ. ομιλία).

#### **Ομοιομορφία ανάλυσης και σχεδίασης**

Τα ΤΝΔ μαθαίνουν και εκτελούν εργασίες με βάση τα δεδομένα που δόθηκαν κατά την εκπαίδευση (ή από αρχική εμπειρία). Ως επεξεργαστές πληροφοριών έχουν έναν κοινό χαρακτήρα που τους επιτρέπει τη χρήση τους σε διαφορετικές εφαρμογές με την εφαρμογή ωστόσο των ίδιων θεωριών και αλγορίθμων μάθησης.

#### **Προσαρμοζόμενη μάθηση**

Τα νευρωνικά δίκτυα μπορούν να επανεκπαιδευτούν για να αντιμετωπίσουν μικρές μεταβολές στις συνθήκες του περιβάλλοντος λειτουργίας τους. Μπορούν ακόμη να σχεδιαστούν κατάλληλα ώστε να λειτουργήσουν και σε ένα μη στατικό περιβάλλον αλλάζοντας τα συναπτικά βάρη σε πραγματικό χρόνο γεγονός που τα καθιστά χρήσιμα εργαλεία για την προσαρμοστική ταξινόμηση προτύπων, την επεξεργασία σήματος και τον προσαρμοστικό έλεγχο.

#### **Ανοχή σφαλμάτων**

Η μερική καταστροφή του δικτύου οδηγεί σε αντίστοιχη μείωση της απόδοσής του. Κάποιες δυνατότητες μπορεί να διατηρηθούν ακόμη και με μεγάλη βλάβη στο δίκτυο. Αυτό συμβαίνει διότι οι νευρώνες λειτουργούν παράλληλα-ταυτόχρονα και ο αριθμός τους μπορεί να είναι πολύ μεγάλος. Αποτελούν χαρακτηριστικό παράδειγμα μαζικά παράλληλου υπολογισμού. Αντίθετα, στους συμβατικούς υπολογιστές για παράδειγμα, η καταστροφή ενός και μόνο στοιχείου μπορεί να οδηγήσει σε ολοκληρωτική κατάρρευση του υπολογιστικού συστήματος [5,7].

## **2.4 Εφαρμογές των τεχνητών νευρωνικών δικτύων**

Τα τεχνητά νευρωνικά δίκτυα ήδη χρησιμοποιούνται για τη λύση διάφορων πολύπλοκων προβλημάτων και αναπτύσσονται συνεχώς για χρήση σε νέες εφαρμογές. Κάποιες από τις περιοχές που τα ΤΝΔ βρίσκουν εφαρμογή σήμερα είναι [6]:

1. Επεξεργασία εικόνας και ηχανική όραση (π.χ. χρήση σε ρομποτικά συστήματα, αναγνώριση χειρόγραφου κειμένου ή φωνής)
2. Επεξεργασία σήματος (π.χ. ανάλυση σεισμικού σήματος)
3. Αναγνώριση προτύπων (π.χ. πιστοποίηση ταυτότητας, ποιοτικός έλεγχος)
4. Αμυντικά συστήματα (π.χ. αναγνώριση στόχου)
5. Οικονομία (π.χ. ανάλυση αγοράς μετοχών, διαχείριση χαρτοφυλακίου, εκτίμηση ακίνητης περιουσίας)
6. Τεχνητή νοημοσύνη (π.χ. υλοποίηση έμπειρων συστημάτων)
7. Εφαρμογές σε PC (π.χ. μηχανές αναζήτησης)

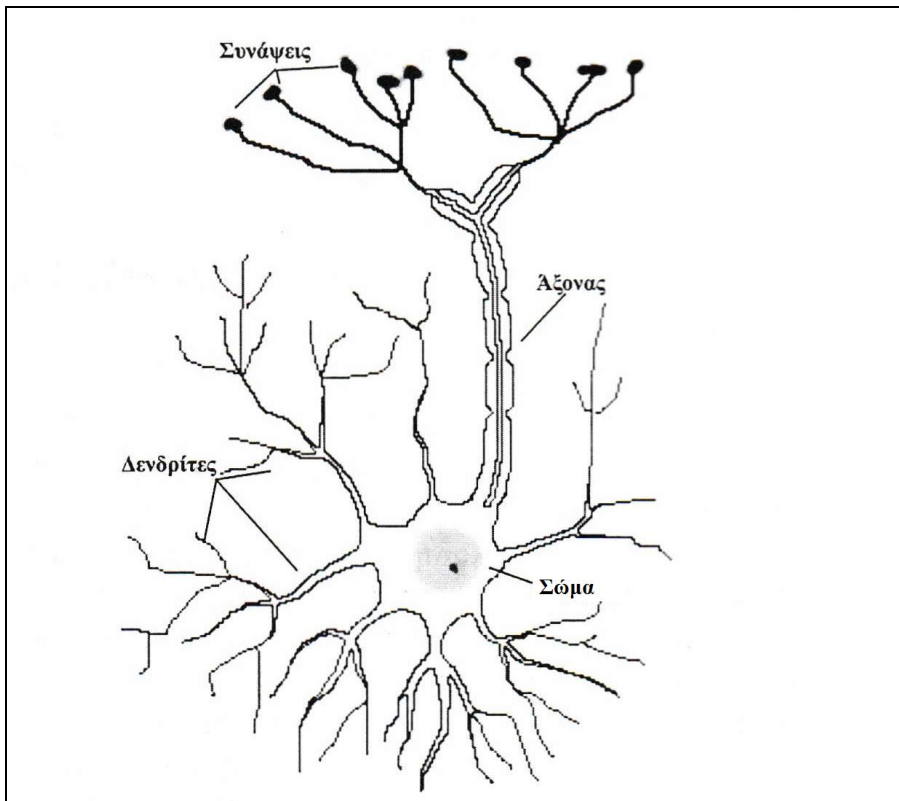
## **2.5 Λειτουργία του βιολογικού νευρωνικού δικτύου**

Η λειτουργία του βιολογικού συστήματος βασίζεται στη διασύνδεση εξειδικευμένων φυσικών κυττάρων που ονομάζονται νευρώνες. Ωστόσο δεν είναι ακόμη επαρκώς γνωστός ο τρόπος με τον οποίον λειτουργεί ο ανθρώπινος εγκέφαλος. Επιπλέον, παρά το γεγονός ότι το βασικό στοιχείο υπολογισμού του ανθρώπινου συστήματος επεξεργασίας είναι σχετικά αργό σε σύγκριση με τους επεξεργαστές, η συνολική επεξεργασία επιτυγχάνεται σε μερικές εκατοντάδες msec. Αυτό μας οδηγεί στο συμπέρασμα ότι η βάση του βιολογικού υπολογισμού είναι ένας μικρός αριθμός ακολουθιακών βημάτων κάθε ένα από τα οποία εκτελείται σε μεγάλο παραλληλισμό [5,6].

Όπως αναφέραμε, το θεμελιώδες δομικό στοιχείο του ανθρώπινου νευρωνικού συστήματος είναι ο νευρώνας ο οποίος σε αναλογία με τους ηλεκτρονικούς υπολογιστές αποτελούν το σύνολο των στοιχείων μνήμης, λογικών κυκλωμάτων και εντολών λειτουργίας. Ένας βιολογικός νευρώνας αποτελείται από:

- τους δενδρίτες
- τις συνάψεις
- το σώμα και
- τον άξονα

Ένας τυπικός νευρώνας συλλέγει σήματα από άλλους νευρώνες μέσω των δενδριτών. Ο νευρώνας στέλνει παλμούς ηλεκτρικής δραστηριότητας μέσω του άξονα ο οποίος διαχωρίζεται στο τέλος του σε πολλά παρακλάδια που ονομάζονται συναπτικές απολήξεις (ή απλά συνάψεις). Η κάθε σύναψη μετατρέπει τη δραστηριότητα από τον άξονα σε ηλεκτρικά αποτελέσματα τα οποία αποτρέπουν ή προκαλούν (διεγείρουν) τη δραστηριότητα στον αντίστοιχο διασυνδεδεμένο νευρώνα [5,6].

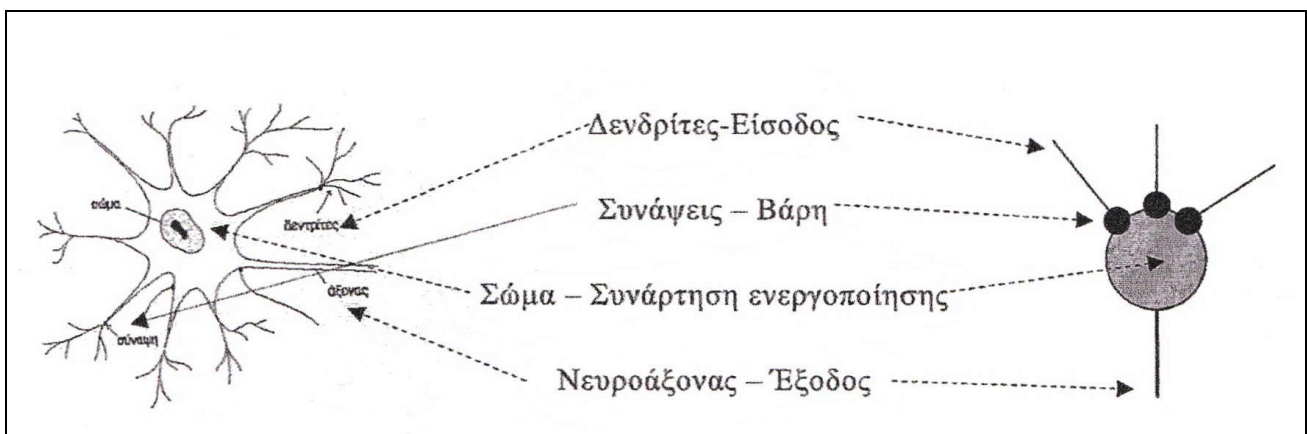


Σχήμα 2.1 Δομή τυπικού βιολογικού νευρώνα [8].

Συνοψίζοντας, λοιπόν, η βασική λειτουργία που επιτελεί ένας νευρώνας είναι:

- Η συσσώρευση των σημάτων που δέχεται από τους νευρώνες με τους οποίους συνδέεται η είσοδός του
- Το φιλτράρισμα και η ενίσχυση αυτών των σημάτων
- Η παραγωγή ενός σήματος εξόδου το οποίο θα μεταδοθεί με όμοιο τρόπο μέσω των συνάψεων προς τους νευρώνες που συνδέονται στην έξοδό του.

Σε πλήρη αντιστοιχία με το απλοποιημένο αυτό μοντέλο του βιολογικού νευρώνα αναπτύχθηκε το μοντέλο του τεχνητού νευρώνα όπως φαίνεται και στο σχήμα 2.2



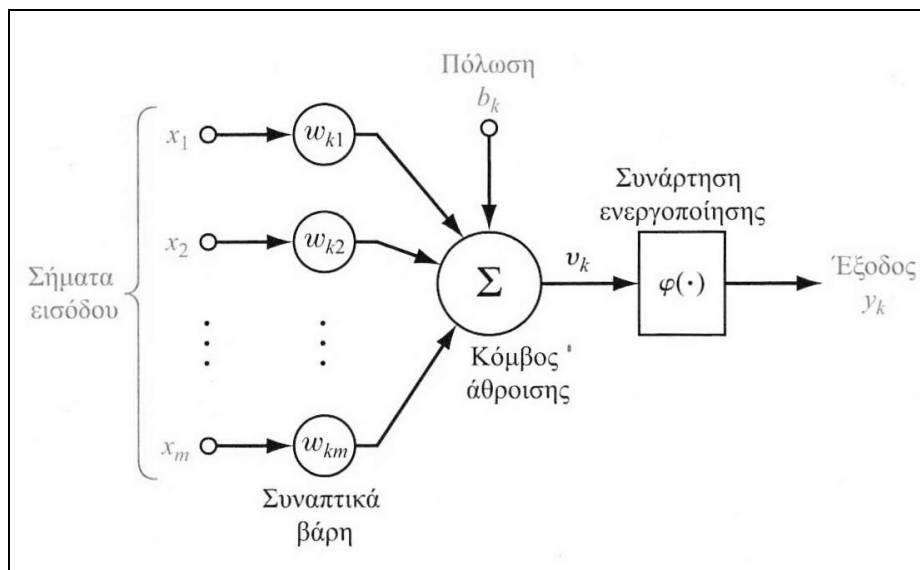
Σχήμα 2.2 Αντιστοιχία βιολογικού νευρώνα με το μοντέλο του τεχνητού νευρώνα [6].

## 2.6 Το μοντέλο του τεχνητού νευρώνα

Το μοντέλο του τεχνητού νευρώνα όπως παρουσιάζεται και παρακάτω στο σχήμα 2.3 αποτελεί τη βάση για τη σχεδίαση μιας μεγάλης οικογένειας νευρωνικών δικτύων. Τα τρία βασικά στοιχεία αυτού του μοντέλου είναι:

- Ένα σύνολο συνάψεων (ή διασυνδέσεων), κάθε μία εκ των οποίων χαρακτηρίζεται από το δικό της βάρος. Συγκεκριμένα, ένα σήμα  $x_j$ , στην είσοδο της συναψης  $j$  που συνδέεται με το νευρώνα  $k$  πολλαπλασιάζεται επί το συναπτικό βάρος  $w_{kj}$ . Ο πρώτος δείκτης στο  $w_{kj}$  αναφέρεται στον εν λόγω νευρώνα και ο δεύτερος δείκτης αναφέρεται στο άκρο εισόδου της συναψης στην οποία αναφέρεται το βάρος. Ανόμοια με το βάρος μιας συναψης στον ανθρώπινο εγκέφαλο, το συναπτικό βάρος ενός τεχνητού νευρώνα μπορεί να λαμβάνει και αρνητικές και θετικές τιμές.
- Έναν αθροιστή (adder) για την άθροιση των σημάτων εισόδου, σταθμισμένων από τα αντίστοιχα συναπτικά βάρη του νευρώνα. Οι λειτουργίες που περιγράφονται εδώ συνιστούν έναν γραμμικό συνδυαστή (linear combiner).
- Μια συνάρτηση ενεργοποίησης (activation function) για τον περιορισμό του πλάτους του σήματος εξόδου ενός νευρώνα. Η συνάρτηση ενεργοποίησης αναφέρεται επίσης και ως συνάρτηση περιορισμού (squashing function) επειδή περιορίζει το επιτρεπτό εύρος πλάτους του σήματος εξόδου σε κάποια πεπερασμένη τιμή (π.χ. 0 έως 1 ή -1 έως 1) [5,6,8].

Το μοντέλο νευρώνα περιλαμβάνει επίσης μια εξωτερικά εφαρμοζόμενη πόλωση, η οποία συμβολίζεται ως  $b_k$ . Η πόλωση αυτή έχει ως αποτέλεσμα την αύξηση ή μείωση της δικτυακής διέγερσης της συνάρτησης ενεργοποίησης, ανάλογα με το εάν είναι θετική ή αρνητική.



Σχήμα 2.3 Μη γραμμικό μοντέλο νευρώνα [5].

Με μαθηματικούς όρους, μπορούμε να περιγράψουμε το νευρώνα  $k$  που απεικονίζεται στο παραπάνω σχήμα 2.3 γράφοντας το ακόλουθο ζεύγος εξισώσεων:

$$v_k = \sum_{j=1}^m w_{kj} x_j \quad (2.1)$$

$$y_k = \varphi(v_k + b_k) \quad (2.2)$$

όπου:

$x_j$  είναι τα σήματα εισόδου ( $j=1,2,\dots,m$ )

$w_{kj}$  είναι τα αντίστοιχα συναπτικά βάρη του  $k$  νευρώνα ( $j=1,2,\dots,m$ )

$u_k$  είναι η έξοδος του γραμμικού συνδυαστή που οφείλεται μόνο στα σήματα εισόδου (δεν παρουσιάζεται στο Σχήμα 2.3)

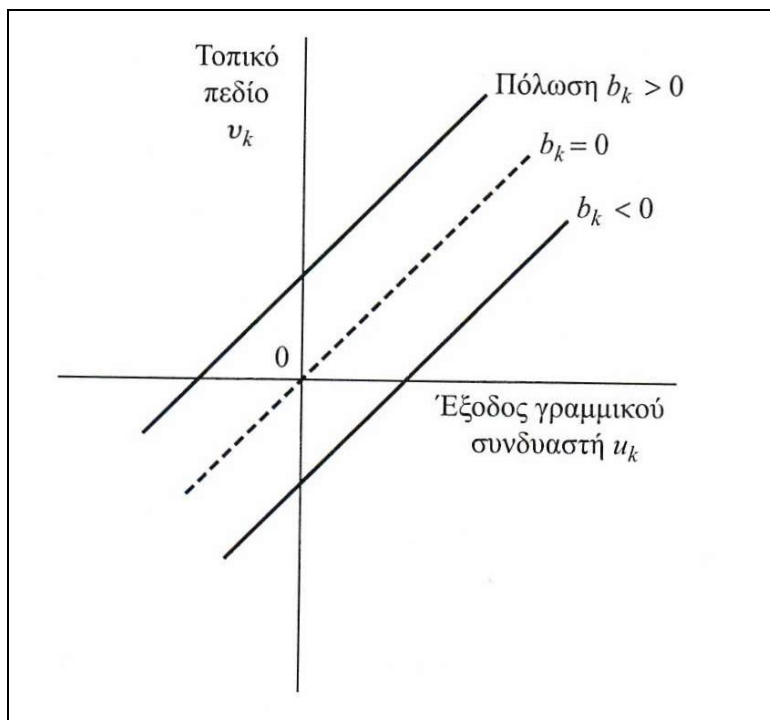
$b_k$  είναι η πόλωση

$\varphi(\cdot)$  είναι η συνάρτηση ενεργοποίησης

$y_k$  είναι το σήμα εξόδου του νευρώνα.

Η χρήση της πόλωσης  $b_k$  έχει ως αποτέλεσμα την εφαρμογή ενός αφινικού μετασχηματισμού στην έξοδο  $u_k$  του γραμμικού συνδυαστή στο μοντέλο του Σχ. 2.3, όπως υποδεικνύει η σχέση:

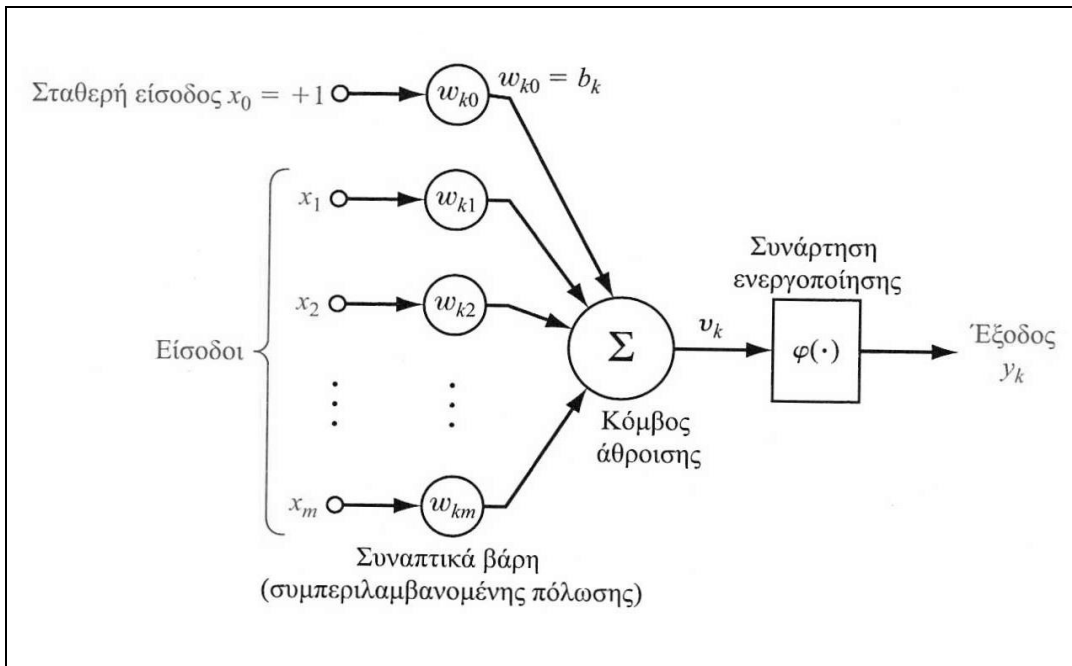
$$v_k = u_k + b_k$$



**Σχήμα 2.4** Αφινικός μετασχηματισμός λόγω ύπαρξης πόλωσης [5].

Ανάλογα με το εάν η πόλωση  $b_k$  είναι θετική ή αρνητική, η σχέση μεταξύ του τοπικού πεδίου, ή δυναμικού ενεργοποίησης  $v_k$  του νευρώνα  $k$  και της εξόδου του γραμμικού συνδυαστή  $u_k$  τροποποιείται όπως φαίνεται στο Σχ. 2.4. Λόγω αυτού του αφινικού μετασχηματισμού, η καμπύλη  $u_k - v_k$  δε διέρχεται πλέον από την αρχή των αξόνων.

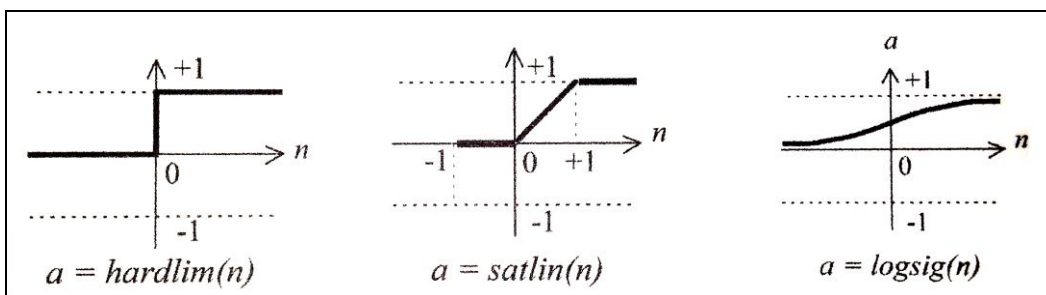
Η πόλωση  $b_k$  είναι μια εξωτερική παράμετρος του νευρώνα  $k$ . Συνυπολογίζουμε την ύπαρξή της προσθέτοντας ένα νέο σήμα εισόδου σταθερό στο  $+1$  και προσθέτοντας ένα νέο συναπτικό βάρος ίσο με την πόλωση  $b_k$  οπότε έχουμε ένα μοντέλο (όπου  $j=0,1,\dots,m$ ) όπως φαίνεται στο σχ. 2.5:



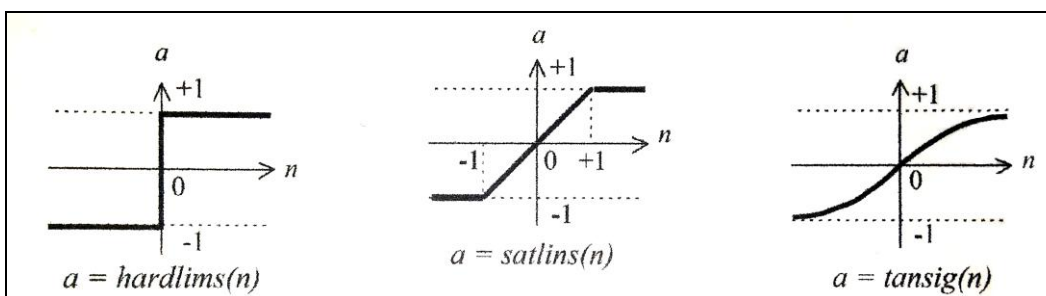
**Σχήμα 2.5** Ένα διαφορετικό μη γραμμικό μοντέλο νευρώνα (το βάρος  $w_{k0}$  αντιστοιχεί στην πόλωση  $b_k$ ) [5].

Η συνάρτηση ενεργοποίησης  $\varphi(\cdot)$  ορίζει την έξοδο ενός νευρώνα βάσει του τοπικού πεδίου  $v_k$ . Η μορφή της συνάρτησης ενεργοποίησης μπορεί να είναι:

- Συνάρτηση λογικής κατωφλίου (σχήμα 2.6)
- Κατά τμήματα γραμμική συνάρτηση (σχήμα 2.6  $\text{satlin}(n)$ )
- Συνάρτηση προσήμου (σχήμα 2.7)
- Υπερβολική εφαπτομένη (σχήμα 2.7  $\text{tansig}(n)$ )



**Σχήμα 2.6** Συναρτήσεις λογικής κατωφλίου [7].



**Σχήμα 2.7** Συναρτήσεις προσήμου [7].



Οι πιο συχνά χρησιμοποιούμενες συναρτήσεις ενεργοποίησης είναι η συνάρτηση λογικής κατωφλίου και η σιγμοειδής συνάρτηση (υπερβολική εφαπτομένη) [5,6,8]. Παρακάτω αναλύουμε τους δύο αυτούς βασικούς τύπους συναρτήσεων ενεργοποίησης:

### 1. Συνάρτηση Κατωφλίου (Threshold Function)

Για τον συγκεκριμένο τύπο συνάρτησης ενεργοποίησης έχουμε:

$$f(v) = \begin{cases} 1 & \text{εάν } v \geq 0 \\ 0 & \text{εάν } v < 0 \end{cases} \quad (2.3)$$

Αντίστοιχα, η έξοδος του νευρώνα  $k$  που χρησιμοποιεί μια τέτοια συνάρτηση κατωφλίου εκφράζεται ως:

$$y_k = \begin{cases} 1 & \text{εάν } v_k \geq 0 \\ 0 & \text{εάν } v_k < 0 \end{cases} \quad (2.4)$$

όπου  $v_k$  είναι το τοπικό πεδίο του νευρώνα, δηλαδή

$$v_k = \sum_{j=1}^m w_{kj} x_j + b_k \quad (2.5)$$

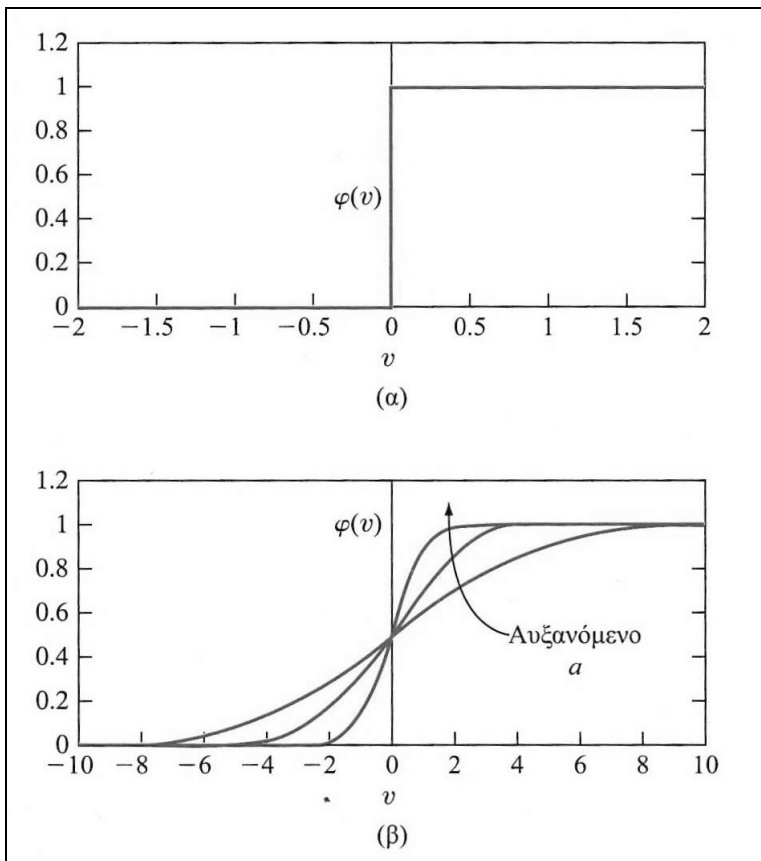
Το παραπάνω μοντέλο αναφέρεται ως μοντέλο McCulloch-Pitts και η έξοδος του λαμβάνει τιμή 1 εάν το τοπικό πεδίο του συγκεκριμένου νευρώνα είναι μη αρνητικό και τιμή 0 σε κάθε άλλη περίπτωση.

### 2. Σιγμοειδής Συνάρτηση (Sigmoid Function)

Η σιγμοειδής συνάρτηση, της οποίας η γραφική παράσταση έχει σχήμα «S», είναι η πλέον κοινή μορφή συνάρτησης ενεργοποίησης που χρησιμοποιείται στην κατασκευή νευρωνικών δικτύων. Ορίζεται ως αυστηρά αύξουσα συνάρτηση και επιδεικνύει ισορροπία μεταξύ γραμμικής και μη γραμμικής συμπεριφοράς [5,6,8]. Ένα παράδειγμα της σιγμοειδούς συνάρτησης είναι η αποκαλούμενη λογιστική συνάρτηση, η οποία ορίζεται ως:

$$f(v) = \frac{1}{1 + e^{-av}} \quad (2.6)$$

όπου  $a$  είναι η παράμετρος κλίσης της σιγμοειδούς συνάρτησης. Μεταβάλλοντας την παράμετρο  $a$ , παίρνουμε σιγμοειδείς συναρτήσεις διαφορετικών κλίσεων, όπως απεικονίζεται στο Σχ. 2.8.



**Σχήμα 2.8** (α) Συνάρτηση καταφλίου και (β) Σιγμοειδής συνάρτηση για μεταβαλλόμενη παράμετρο  $a$  [5].

Η κλίση στο σημείο αρχής των αξόνων ισούται με  $a/4$ . Στο όριο, καθώς η παράμετρος κλίσης προσεγγίζει το άπειρο, η σιγμοειδής συνάρτηση γίνεται απλώς μια συνάρτηση καταφλίου. Ενώ μια συνάρτηση καταφλίου λαμβάνει τιμή 0 ή 1, μια σιγμοειδής συνάρτηση μπορεί να λαμβάνει τιμές από ένα συνεχές πεδίο τιμών, από 0 έως 1. Επίσης, η σιγμοειδής συνάρτηση είναι διαφορίσιμη, ενώ η συνάρτηση καταφλίου όχι (η διαφορισιμότητα είναι ένα σημαντικό χαρακτηριστικό της θεωρίας των νευρωνικών δικτύων). Οι συναρτήσεις ενεργοποίησης που ορίζουν οι Εξ.2.3 και 2.6 έχουν πεδίο τιμών από 0 έως +1. Σε ορισμένες περιπτώσεις, είναι επιθυμητό να έχουμε για τη συνάρτηση ενεργοποίησης πεδίο τιμών από -1 έως +1. Τότε, η συνάρτηση ενεργοποίησης είναι μια περιττή συνάρτηση του τοπικού πεδίου [5,6,8]. Η συνάρτηση καταφλίου της Εξ.2.3 ορίζεται τώρα ως:

$$f(v) = \begin{cases} +1 & \text{εάν } v > 0 \\ 0 & \text{εάν } v = 0 \\ -1 & \text{εάν } v < 0 \end{cases} \quad (2.7)$$

η οποία είναι γνωστή ως συνάρτηση προσήμου. Για την αντίστοιχη μορφή μιας σιγμοειδούς συνάρτησης, μπορούμε να χρησιμοποιήσουμε τη συνάρτηση υπερβολικής εφαπτομένης η οποία ορίζεται ως:

$$f(v) = \tanh(v) \quad (2.8)$$

Το να επιτρέπουμε σε μια συνάρτηση ενεργοποίησης του σιγμοειδούς τύπου να λαμβάνει αρνητικές τιμές όπως η Εξ.2.8 μπορεί να μας παρέχει πρακτικά οφέλη έναντι της λογιστικής συνάρτησης της Εξ.2.6.

## **2.7 Τα τεχνητά νευρωνικά δίκτυα (ΤΝΔ)**

Με τη διασύνδεση πολλών τεχνητών νευρώνων επιτυγχάνουμε τη δημιουργία του τεχνητού νευρωνικού δικτύου (ΤΑΝ). Τα ΤΝΔ να αποτελούν μια προσεγγιστική αναπαράσταση του πραγματικού συστήματος. Κατά την προσομοίωση αυτό που κάνουμε είναι να προγραμματίζουμε έναν υπολογιστή με τη βοήθεια του μαθηματικού μοντέλου που αναπαριστά το νευρωνικό δίκτυο ώστε να "αντιλαμβάνεται" ο υπολογιστής τα χαρακτηριστικά των νευρώνων και των διασυνδέσεων μεταξύ τους. Ένα νευρωνικό δίκτυο ορίζεται ως ένας προσανατολισμένος γράφος όπου:

- Εισάγεται σε κάθε κόμβο  $k$  μια μεταβλητή κατάστασης  $x_j$
- Για κάθε διασύνδεση μεταξύ κόμβων ανατίθεται ένα βάρος  $w_{kj}$
- Για κάθε κόμβο  $k$  δίνεται μια τιμή πόλωσης  $b_k$
- Για κάθε κόμβο  $k$  ορίζεται μια συνάρτηση ενεργοποίησης  $f(x_j, w_{kj}, b_k)$  η οποία καθορίζει την κατάσταση του κόμβου συναρτήσει της κατάστασης των κόμβων που στέλνουν το σήμα τους στον κόμβο  $k$ , των βαρών των εισερχόμενων αυτών συνδέσεων και της πόλωσής του

Στην ορολογία των ΤΝΔ οι κόμβοι ονομάζονται νευρώνες (neurons), οι συνδέσμοι ονομάζονται συνάψεις (synapses) και φέρουν τα αντίστοιχα συναπτικά βάρη (weights), η πόλωση (bias) είναι γνωστή ως κατώφλι ενεργοποίησης (activation threshold) [5,6,8].

Κάθε νευρώνας εκτελεί μια σχετικά απλή διεργασία: δέχεται τις εισόδους των γειτονικών νευρώνων ή τις εξωτερικές εισόδους και τις χρησιμοποιεί για να υπολογίσει το σήμα εξόδου το οποίο διαδίδεται στους άλλους νευρώνες. Επίσης, κάθε νευρώνας εκτελεί και την προσαρμογή των βαρών. Το σύστημα είναι από τη φύση του παράλληλο με την έννοια ότι πολλοί κόμβοι μπορούν να εκτελούν τους υπολογισμούς τους ταυτόχρονα. Παρακάτω διαχωρίζουμε τρία είδη κόμβων:

- Οι κόμβοι χωρίς εισερχόμενους συνδέσμους ονομάζονται νευρώνες εισόδου (input neurons) και δέχονται δεδομένα έξω από το ΝΔ.
- Οι κόμβοι χωρίς εξερχόμενους συνδέσμους ονομάζονται νευρώνες εξόδου (output neurons) και στέλνουν δεδομένα έξω από το ΝΔ.
- Οι κόμβοι με εισερχόμενους και εξερχόμενους συνδέσμους ονομάζονται κρυμμένοι νευρώνες (hidden neurons). Οι εισοδοί και εξοδοί τους παραμένουν μέσα στο ΤΝΔ.  
[5,6,8]

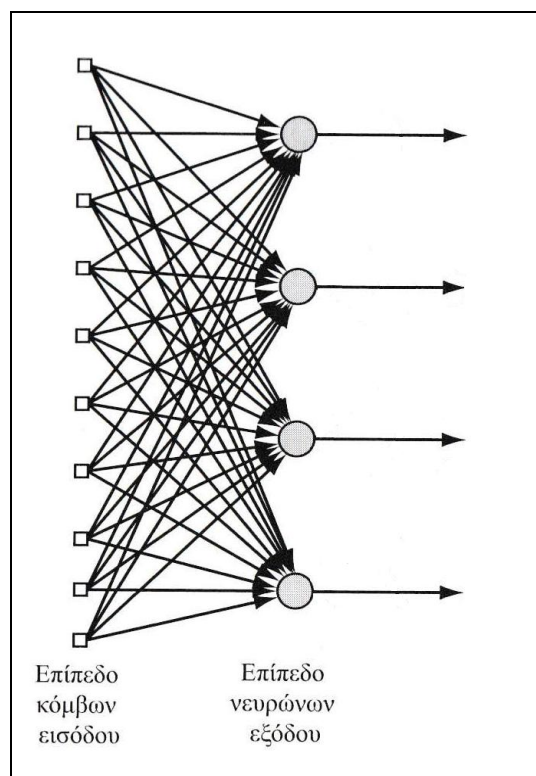
## **2.8 Αρχιτεκτονικές τεχνητών νευρωνικών δικτύων**

Ο τρόπος με τον οποίο είναι δομημένοι οι νευρώνες ενός νευρωνικού δικτύου σχετίζεται στενά με τον αλγόριθμο μάθησης που χρησιμοποιείται για την εκπαίδευση του δικτύου. Άρα, μπορούμε να αντιμετωπίσουμε τους αλγόριθμους μάθησης (κανόνες) που χρησιμοποιούνται στη σχεδίαση νευρωνικών δικτύων ως δομημένους. Γενικά, μπορούμε να αναγνωρίζουμε τρεις θεμελιωδώς διαφορετικές κατηγορίες αρχιτεκτονικών δικτύων:

### **1. Ενός Επιπέδου Δίκτυα Πρόσθιας Τροφοδότησης**

Σε ένα νευρωνικό δίκτυο, οι νευρώνες οργανώνονται σε μορφή επιπέδων. Στην απλούστερη δυνατή μορφή ενός δικτύου, έχουμε ένα επίπεδο εισόδου το οποίο συνδέεται απευθείας με ένα επίπεδο νευρώνων εξόδου,

αλλά όχι αντίστροφα. Αυτό το δίκτυο είναι τύπου πρόσθιας τροφοδότησης (feedforward). Απεικονίζεται στο Σχ. 2.9 για την περίπτωση δέκα κόμβων στο επίπεδο εισόδου και τεσσάρων στο επίπεδο εξόδου. Ένα τέτοιο δίκτυο αποκαλείται δίκτυο ενός επιπέδου, με το χαρακτηρισμό «ενός επιπέδου» να αναφέρεται στο επίπεδο εξόδου, των υπολογιστικών κόμβων. Το επίπεδο εισόδου δεν προσμετράται επειδή δεν εκτελείται κανένας υπολογισμός σε αυτό [5,6,8].

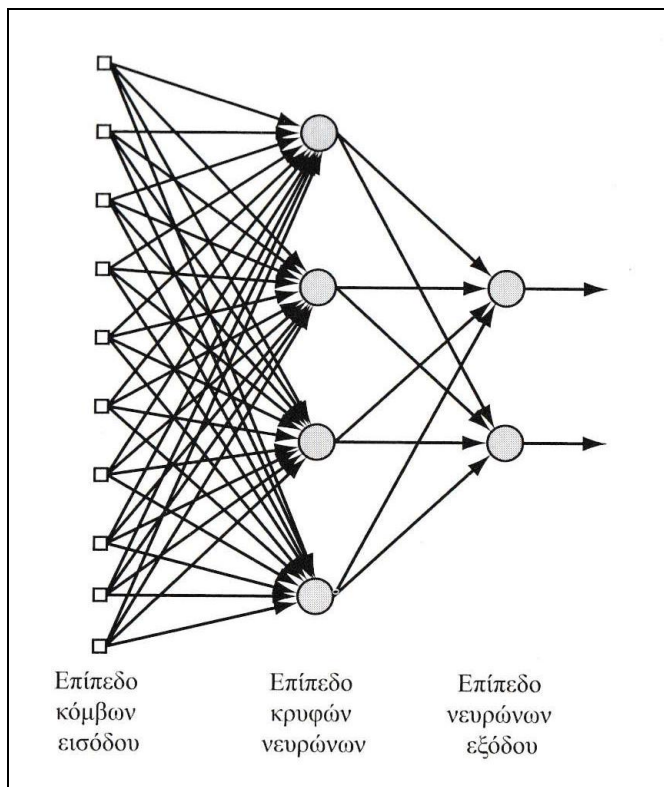


**Σχήμα 2.9** Δίκτυο πρόσθιας τροφοδότησης με ένα μεμονωμένο επίπεδο νευρώνων [5].

## 2. Πολυεπίπεδα Δίκτυα Πρόσθιας Τροφοδότησης

Η δεύτερη κατηγορία νευρωνικών δικτύων πρόσθιας τροφοδότησης χαρακτηρίζεται από την παρουσία ενός ή περισσότερων κρυφών επιπέδων, των οποίων οι υπολογιστικοί κόμβοι αποκαλούνται κρυφοί νευρώνες ή κρυφές μονάδες. Ο όρος «κρυφός» αναφέρεται στο γεγονός ότι αυτό το μέρος του νευρωνικού δικτύου δεν είναι άμεσα ορατό ούτε από την είσοδο, ούτε από την έξοδο του δικτύου. Η λειτουργία των κρυφών νευρώνων είναι να παρεμβαίνουν μεταξύ της εξωτερικά προερχόμενης εισόδου και της εξόδου του δικτύου με κάποιο χρήσιμο τρόπο. Προσθέτοντας ένα ή περισσότερα κρυφά επίπεδα, το δίκτυο έχει τη δυνατότητα να εξαγάγει στατιστικά αποτελέσματα υψηλότερης τάξης από την είσοδό του. Οι κόμβοι στο επίπεδο εισόδου του δικτύου παρέχουν τα αντίστοιχα στοιχεία του προτύπου ενεργοποίησης (διάνυσμα εισόδου), τα οποία συνιστούν τα σήματα εισόδου που εφαρμόζονται στους νευρώνες (υπολογιστικούς κόμβους) του δεύτερου επιπέδου (δηλ. του πρώτου κρυφού επιπέδου). Τα σήματα εξόδου του δεύτερου επιπέδου χρησιμοποιούνται ως εισοδοί στο τρίτο επίπεδο, και αυτό συνεχίζεται για όλα τα υπόλοιπα επίπεδα του δικτύου. Τυπικά, οι νευρώνες σε κάθε επίπεδο του δικτύου έχουν ως εισόδους τα σήματα εξόδου μόνο του προηγούμενου επιπέδου. Το σύνολο των σημάτων εξόδου των νευρώνων στο επίπεδο εξόδου (το τελικό επίπεδο) του δικτύου συνιστά την συνολική απόκριση του δικτύου στο πρότυπο ενεργοποίησης που παρέχεται από τους πηγαίους κόμβους στο επίπεδο εισόδου (πρώτο επίπεδο του δικτύου). Στο Σχ.2.10 απεικονίζεται η διάταξη ενός πολυεπίπεδου νευρωνικού δικτύου πρόσθιας τροφοδότησης για την περίπτωση ενός και μόνο κρυφού επιπέδου. Το δίκτυο του Σχ.2.10 αναφέρεται ως δίκτυο 10-4-2 επειδή έχει 10 κόμβους εισόδου, 4 κρυφούς

νευρώνες και 2 νευρώνες εξόδου. Γενικά, ένα δίκτυο πρόσθιας τροφοδότησης με  $m$  πηγαίους κόμβους,  $h_1$  νευρώνες στο πρώτο κρυφό επίπεδο,  $h_2$  νευρώνες στο δεύτερο κρυφό επίπεδο και  $q$  νευρώνες στο επίπεδο εξόδου, αναφέρεται ως δίκτυο  $m-h_1-h_2-q$  [5].

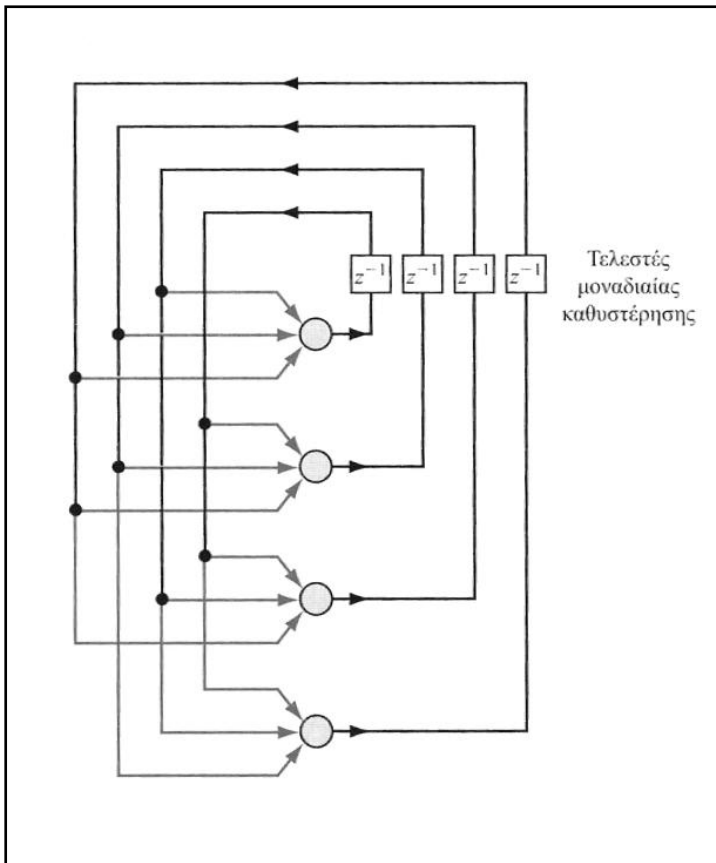


**Σχήμα 2.10** Πλήρως συνδεδεμένο δίκτυο πρόσθιας τροφοδότησης με ένα κρυφό επίπεδο και ένα επίπεδο εξόδου [5].

Το νευρωνικό δίκτυο του Σχ.2.10 λέγεται "πλήρως συνδεδεμένο" αφού κάθε κόμβος σε κάθε επίπεδο του δικτύου συνδέεται με κάθε άλλο κόμβο του επόμενου (προς τα εμπρός) επιπέδου διαφορετικά αναφέρεται ως μερικώς συνδεδεμένο.

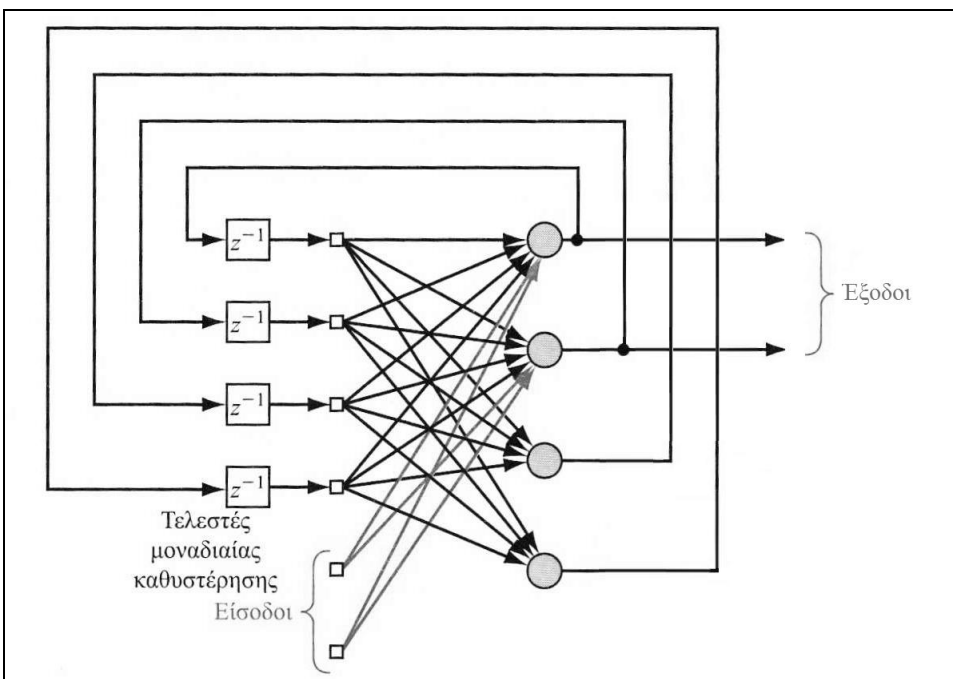
### 3. Αναδρομικά Δίκτυα

Ένα αναδρομικό νευρωνικό δίκτυο (Recurrent Neural Network) διαφέρει από ένα νευρωνικό δίκτυο πρόσθιας τροφοδότησης στο ότι έχει τουλάχιστον ένα βρόχο ανάδρασης. Για παράδειγμα, ένα αναδρομικό δίκτυο μπορεί να αποτελείται από ένα μεμονωμένο επίπεδο νευρώνων, με κάθε νευρώνα να τροφοδοτεί το σήμα εξόδου του πίσω, στις εισόδους όλων των άλλων νευρώνων, όπως φαίνεται στο Σχ.2.11. Το αναδρομικό δίκτυο του Σχ.2.11 δεν έχει κρυφούς νευρώνες [5].



Σχήμα 2.11 Αναδρομικό δίκτυο χωρίς βρόχους αυτο-ανάδρασης και κρυφούς νευρώνες [5].

Μια άλλη κατηγορία αναδρομικών δικτύων με κρυφούς νευρώνες φαίνεται στο Σχ.2.12. Οι συνδέσεις ανάδρασης προέρχονται από τους κρυφούς νευρώνες καθώς και από τους νευρώνες εξόδου.



Σχήμα 2.12 Αναδρομικό δίκτυο με κρυφούς νευρώνες [5].

Οι βρόχοι ανάδρασης προϋποθέτουν τη χρήση συγκεκριμένων κλάδων, αποτελούμενων από στοιχεία μοναδιαίας χρονικής καθυστέρησης (συμβολίζονται ως  $z^{-1}$ ), τα οποία έχουν ως αποτέλεσμα μη γραμμική δυναμική συμπεριφορά, εάν υποθέσουμε ότι το νευρωνικό δίκτυο περιέχει μη γραμμικές μονάδες [5].

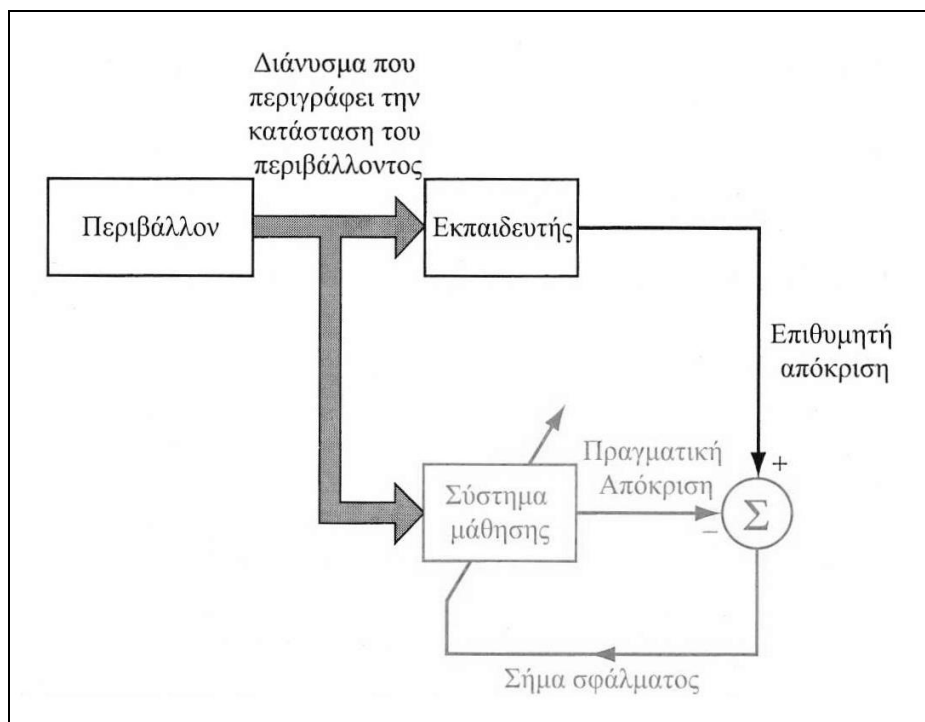
## 2.9 Διαδικασίες εκπαίδευσης ΤΝΔ

Στα ΝΔ η μάθηση αναφέρεται στη διεργασία επίτευξης μιας επιθυμητής συμπεριφοράς μέσω ανανέωσης των τιμών των συναπτικών βαρών, που τους επιτρέπει να μαθαίνουν από το περιβάλλον τους και να βελτιώνουν τη συμπεριφορά τους με το πέρασμα του χρόνου. Υπάρχουν διάφορες μέθοδοι προσαρμογής των βαρών των συνδέσεων, που συνιστούν διαφορετικούς αλγόριθμους μάθησης ή εκπαίδευσης. Κάθε αλγόριθμος μάθησης αποτελεί ένα προκαθορισμένο σύνολο κανόνων επίλυσης του προβλήματος μάθησης και έχει συγκεκριμένα πλεονεκτήματα και μειονεκτήματα [5].

## ΜΕΘΟΔΟΙ ΕΚΠΑΙΔΕΥΣΗΣ

### 1. Μάθηση Με Εκπαιδευτή

Η μάθηση με εκπαιδευτή αναφέρεται επίσης ως επιβλεπόμενη μάθηση (βλ.Σχ.2.13). Μπορούμε να θεωρήσουμε ότι ο εκπαιδευτής έχει γνώση του περιβάλλοντος, και αυτή η γνώση αντιπροσωπεύεται από ένα σύνολο παραδειγμάτων εισόδου-εξόδου. Ωστόσο, το περιβάλλον είναι άγνωστο στο νευρωνικό δίκτυο. Ας υποθέσουμε τώρα ότι ο εκπαιδευτής και το νευρωνικό δίκτυο εκτίθενται σε ένα διάνυσμα εκπαίδευσης (δηλ., ένα παράδειγμα) αντλούμενο από το ίδιο περιβάλλον. Ο εκπαιδευτής είναι σε θέση να παρέχει στο νευρωνικό δίκτυο μια επιθυμητή απόκριση για το συγκεκριμένο διάνυσμα εκπαίδευσης [5,8].



**Σχήμα 2.13** Διάγραμμα της μάθησης με εκπαιδευτή. Το έγχρωμο μέρος του σχήματος αποτελεί ένα βρόχο ανάδρασης [5].

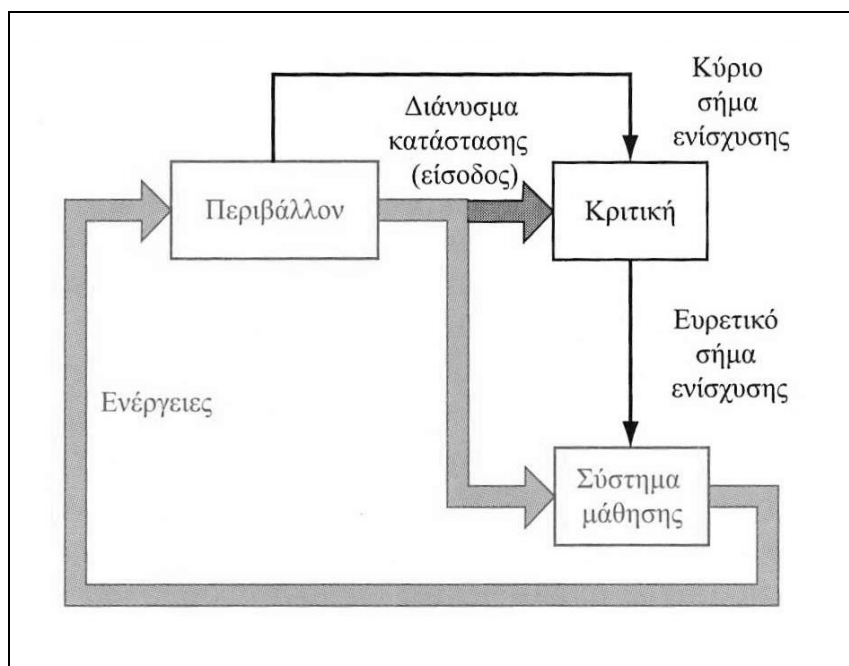
Η επιθυμητή απόκριση αντιπροσωπεύει τη «βέλτιστη» ενέργεια που πρέπει να εκτελείται από το νευρωνικό δίκτυο. Οι παράμετροι του δικτύου προσαρμόζονται υπό τη συνδυασμένη επιρροή του διανύσματος εκπαίδευσης και του σήματος σφάλματος. Το σήμα σφάλματος ορίζεται ως η διαφορά μεταξύ της επιθυμητής απόκρισης και της πραγματικής απόκρισης του δικτύου. Αυτή η προσαρμογή εκτελείται με επαναληπτικό τρόπο, βήμα προς βήμα, με στόχο να φέρει τελικά το νευρωνικό δίκτυο σε μια κατάσταση όπου θα προσομοιώνει τη συμπεριφορά του εκπαιδευτή. Έτσι, η γνώση του περιβάλλοντος που είναι διαθέσιμη στον εκπαιδευτή μεταφέρεται στο νευρωνικό δίκτυο μέσω εκπαίδευσης και αποθηκεύεται με τη μορφή «σταθερών» συναπτικών βαρών. Όταν επιτευχθεί αυτή η συνθήκη, μπορούμε να απαλλαγούμε από τον εκπαιδευτή και να αφήσουμε το νευρωνικό δίκτυο να αντιμετωπίσει το περιβάλλον εντελώς μόνο του [5].

## **2. Μάθηση Χωρίς Εκπαιδευτή**

Στην επιβλεπόμενη μάθηση, η διαδικασία μάθησης λαμβάνει χώρα υπό την καθοδήγηση ενός εκπαιδευτή. Στη μάθηση χωρίς εκπαιδευτή, δεν υπάρχει εκπαιδευτής που να επιβλέπει τη διαδικασία μάθησης. Δηλαδή, δεν υπάρχουν χαρακτηρισμένα παραδείγματα της λειτουργίας που πρέπει να μάθει το δίκτυο. Μπορούμε να ορίσουμε δύο υποκατηγορίες μάθησης:

### **A. Ενισχυτική Μάθηση**

Στην ενισχυτική μάθηση, η εκμάθηση μιας αντιστοίχισης εισόδου-εξόδου εκτελείται μέσω συνεχούς αλληλεπίδρασης με το περιβάλλον, με στόχο την ελαχιστοποίηση ενός βαθμωτού δείκτη απόδοσης. Το Σχ.2.14 παρουσιάζει το σχηματικό διάγραμμα μιας μορφής ενός συστήματος ενισχυτικής μάθησης που βασίζεται σε ένα μηχανισμό που λειτουργεί ως κριτής, ο οποίος μετατρέπει ένα κύριο σήμα ενίσχυσης λαμβανόμενο από το περιβάλλον σε ένα υψηλότερης ποιότητας σήμα ενίσχυσης που αποκαλείται ευρετικό σήμα ενίσχυσης. Ο στόχος της ενισχυτικής μάθησης είναι να ελαχιστοποιεί μια συνάρτηση τρέχοντος κόστους, η οποία ορίζεται ως πρόβλεψη του αθροιστικού κόστους ενεργειών που εκτελούνται σε μια αλληλουχία βημάτων αντί απλώς του άμεσου κόστους μιας ενέργειας [5,8].



**Σχήμα 2.14** Σχηματικό διάγραμμα της ενισχυτικής μάθησης. Το σύστημα μάθησης και το περιβάλλον είναι αμφότερα εντός του βρόχου ανάδρασης [5].



## B. Μη Επιβλεπόμενη Μάθηση

Στη μη επιβλεπόμενη μάθηση δεν υπάρχει εξωτερικός εκπαιδευτής ή κριτής που να επιβλέπει τη διαδικασία μάθησης. Αντίθετα, υπάρχει ένα ανεξάρτητο από την εργασία μέτρο της ποιότητας της αναπαράστασης που καλείται να μάθει το δίκτυο και οι ελεύθερες παράμετροι του δικτύου βελτιστοποιούνται σε σχέση με αυτό το μέτρο. Έτσι, αναπτύσσει τη δυνατότητα να σχηματίζει εσωτερικές αναπαραστάσεις για την κωδικοποίηση χαρακτηριστικών της εισόδου και να δημιουργεί νέες κλάσεις αυτόματα.

Εδώ μπορούμε να χρησιμοποιήσουμε έναν κανόνα ανταγωνιστικής μάθησης δηλαδή, π.χ. να χρησιμοποιήσουμε ένα νευρωνικό δίκτυο το οποίο αποτελείται από δύο επίπεδα (ένα επίπεδο εισόδου και ένα ανταγωνιστικό επίπεδο). Το επίπεδο εισόδου λαμβάνει τα διαθέσιμα δεδομένα. Το ανταγωνιστικό επίπεδο αποτελείται από νευρώνες οι οποίοι ανταγωνίζονται ο ένας τον άλλο (σύμφωνα με έναν κανόνα μάθησης) για την «ευκαιρία» να αποκριθούν σε χαρακτηριστικά που περιέχονται στα δεδομένα εισόδου. Στην απλούστερη δυνατή μορφή του, το δίκτυο λειτουργεί σύμφωνα με μια στρατηγική «ο νικητής παίρνει τα πάντα». Βάσει μιας τέτοιας στρατηγικής, ο νευρώνας με τη μεγαλύτερη συνολική είσοδο «νικά» στον ανταγωνισμό και ενεργοποιείται ενώ όλοι οι άλλοι νευρώνες του δικτύου απενεργοποιούνται [5,8].

## ΒΑΣΙΚΟΙ ΤΡΟΠΟΙ ΝΕΥΡΩΝΙΚΗΣ ΜΑΘΗΣΗΣ

### 1. Μάθηση Διόρθωσης Σφάλματος

Σε αυτόν τον τρόπο μάθησης το ΤΝΔ προσαρμόζει τα συναπτικά βάρη συγκρίνοντας τις προκύπτουσες εξόδους  $y_k(t)$  (για συγκεκριμένες εισόδους  $x_k(t)$ ) με τις επιθυμητές εξόδους  $d_k(t)$ . Με βάση το σφάλμα :

$$e_k(t) = d_k(t) - y_k(t) \quad (2.9)$$

ορίζεται το κριτήριο :

$$I = E \left\{ \frac{1}{2} \sum_k e_k^2(t) \right\} \quad (2.10)$$

το οποίο καλείται κριτήριο μέσου τετραγωνικού σφάλματος και εκφράζει τη μέση τιμή του αθροίσματος των τετραγώνων των σφαλμάτων. Πλέον, το πρόβλημα μάθησης είναι να επιλεγούν τα συναπτικά βάρη των νευρώνων ώστε να ελαχιστοποιηθεί το μέσο τετραγωνικό σφάλμα [5,7].

Το στιγμιαίο κριτήριο τετραγωνικού σφάλματος είναι :

$$J = \frac{1}{2} \sum_k e_k^2(t) \quad (2.11)$$

και ο κανόνας μάθησης γίνεται :

$$\Delta w_{kj}(t) = \gamma e_k(t) x_k(t) \quad (2.12) \quad (\text{κανόνας δέλτα ή κανόνας Widrow-Hoff})$$

όπου :  $w_{kj}$  είναι το βάρος της σύναψης  $j$  του νευρώνα  $k$

$\Delta w_{kj}$  είναι η ποσότητα αλλαγής του βάρους στο χρόνο  $t$

και  $\gamma$  είναι μια θετική σταθερά που καθορίζει το ρυθμό μάθησης (learning rate)

Η νέα τιμή του θεωρούμενου συναπτικού βάρους είναι :

$$w_{kj}(t+1) = w_{kj}(t) + \Delta w_{kj}(t) \quad (2.13)$$

Η βήμα προς βήμα μεταβολή των βαρών συνεχίζεται μέχρι το ΤΝΔ να φτάσει σε κατάσταση ισορροπίας. Ο αλγόριθμος μάθησης μέσω διόρθωσης σφαλμάτων έχει τοπική φύση και η σύγκλισή του εξαρτάται από την επιλογή του ρυθμού εκμάθησης  $\gamma$ .

## **2. Μάθηση Hebb**

Στη μάθηση αυτού του τύπου :

- Αν δύο νευρώνες που συνδέονται με μια σύναψη ενεργοποιούνται ταυτόχρονα τότε το βάρος της σύναψης αυτής ενισχύεται (αυξάνεται).
- Αν δύο νευρώνες που συνδέονται με μια σύναψη ενεργοποιούνται ασύγχρονα τότε το βάρος της σύναψης αυτής εξασθενεί ( μειώνεται) ή εξαλείφεται επιλεκτικά.

Μια τέτοια σύναψη ονομάζεται Εββιανή (τύπου Hebb) και αυξάνει τη συνοπτική αποδοτικότητα ως συνάρτηση της συσχέτισης ανάμεσα στις προσυναπτικές και μετασυναπτικές δραστηριότητες. Αυτό γίνεται με τη χρήση ενός χρονικά εξαρτώμενου και ισχυρά τοπικού μηχανισμού [5,7].

Το μαθηματικό μοντέλο της τροποποίησης των βαρών σε μια σύναψη Hebb είναι της μορφής :

$$\Delta w_{kj}(t) = H(y_k(t)x_k(t)) \quad (2.14)$$

όπου :  $H()$  είναι μια συνάρτηση της προσυναπτικής και μετασυναπτικής δραστηριότητας ή σε απλούστερη μορφή :

$$\Delta w_{kj}(t) = \gamma y_k(t)x_k(t) \quad (2.15)$$

όπου  $\gamma$  είναι ο ρυθμός μάθησης.

## **3. Ανταγωνιστική μάθηση**

Σε αυτού του είδους τη μάθηση οι νευρώνες εξόδου ανταγωνίζονται μεταξύ τους για το ποιος θα ενεργοποιηθεί και κάθε χρονική στιγμή μόνο ένας νευρώνας μπορεί να είναι ενεργός, σε αντίθεση με την Εββιανή μάθηση (winner takes all). Σε αυτή την ιδιότητα οφείλεται η ικανότητα της ανταγωνιστικής μάθησης να εξάγει ιδιότητες και να ταξινομεί εισόδους επιτυχώς. Ο νευρώνας που κερδίζει ονομάζεται νευρώνας-νικητής. Οι νευρώνες είναι όλοι ισοδύναμοι, διαφέρουν μόνο στα βάρη, και κάθε νευρώνας "ειδικεύεται" στη μάθηση συνόλων ομοειδών εισόδων οπότε και εξελίσσεται σε ανιχνευτή χαρακτηριστικών. Θεωρούμε ένα μονοστρωματικό ΤΝΔ με εξόδους πλήρως συνδεδεμένες με τις εισόδους και παράπλευρες συνδέσεις που δρουν απαγορευτικά [5,7].

Νικητής νευρώνας είναι αυτός με το μεγαλύτερο δυναμικό ενεργοποίησης. Η έξοδος του νευρώνα-νικητή τοποθετείται στην τιμή 1 ενώ των υπόλοιπων νευρώνων στην τιμή 0. Κάθε νευρώνας λαμβάνει μια σταθερή ποσότητα συναπτικού βάρους το οποίο κατανέμεται στους κόμβους εισόδου ώστε να ισχύει:

$$\sum_j w_{jk} = 1 \text{ για όλα τα } k \quad (2.16)$$

Υπάρχει δηλαδή περιορισμός στην ισχύ (άθροισμα βαρών) των νευρώνων.

Ο απλούστερος κανόνας ανταγωνιστικής μάθησης είναι ο εξής :

$$\Delta w_{kj}(t) = \begin{cases} \gamma(x_j - w_{kj}) & \text{αν ο } k \text{ νευρώνας κερδίζει} \\ 0 & \text{αν ο } k \text{ νευρώνας χάνει} \end{cases} \quad (2.17)$$

Προφανώς, μέσω αυτού του κανόνα το διάνυσμα των βαρών του νικητή μετακινείται προς το διάνυσμα εισόδου.

#### **4. Μάθηση Boltzmann**

Η μάθηση αυτή που αναπτύχθηκε από τον Boltzmann είναι μια στοχαστική (πιθανοτική) μάθηση βασισμένη στη θερμοδυναμική. Χρησιμοποιείται στη μηχανή Boltzmann που είναι κατά βάση ένα αναδρομικό ΤΝΔ, οι νευρώνες του οποίου λειτουργούν δυαδικά, δηλαδή έχουν δύο επιτρεπτές καταστάσεις, την κατάσταση ON (+1) και την κατάσταση OFF (-1).

Η μηχανή Boltzmann λειτουργεί με δύο διακριτούς τρόπους :

- Υπό συνθήκες σύσφιξης
- Υπό συνθήκες ελεύθερης λειτουργίας

Στον πρώτο τρόπο λειτουργίας όλοι οι ορατοί νευρώνες (αυτοί δηλαδή που παρέχουν επικοινωνία μεταξύ του ΤΝΔ και του περιβάλλοντος) συσφίγγονται σε συγκεκριμένες καταστάσεις που καθορίζονται από το περιβάλλον. Στο δεύτερο τρόπο και οι ορατοί και οι κρυμμένοι νευρώνες (δηλαδή αυτοί που λειτουργούν πάντα ελεύθερα) επιτρέπεται να λειτουργούν ελεύθερα [5,7].

## **2.10 Εργασίες μάθησης**

Η επιλογή ενός συγκεκριμένου κανόνα μάθησης επηρεάζεται από τις εργασίες μάθησης, η ποικιλότητα φύση των οποίων καταδεικνύει την καθολικότητα εφαρμογής των νευρωνικών δικτύων.

### **A. Συσχέτιση Προτύπων**

Η συσχέτιση μπορεί να λάβει μία από τις εξής δύο μορφές:

- Αυτο-συσχέτιση, όπου ένα νευρωνικό δίκτυο καλείται να αποθηκεύσει ένα σύνολο προτύπων (διάνυσματα), τα οποία παρουσιάζονται κατ' επανάληψη στο δίκτυο. Στη συνέχεια παρουσιάζεται στο δίκτυο μια ημιτελής περιγραφή ή παραμορφωμένη (με θόρυβο) έκδοση ενός εκ των αρχικών προτύπων που έχει αποθηκεύσει και το ζητούμενο είναι να ανακτήσει (να ανακαλέσει) αυτό το συγκεκριμένο πρότυπο από τη μνήμη του.
- Ετερο-συσχέτιση, η οποία διαφέρει από την αυτο-συσχέτιση στο ότι ένα τυχαίο σύνολο προτύπων εισόδου συνδυάζεται με ένα άλλο τυχαίο σύνολο προτύπων εξόδου.

Η αυτο-συσχέτιση προϋποθέτει τη χρήση μη επιβλεπόμενης μάθησης, ενώ στην έτερο-συσχέτιση η μάθηση είναι επιβλεπόμενη [5,7].

Έστω ότι το  $x_k$  είναι ένα πρότυπο (διάνυσμα) που εφαρμόζεται σε μια συσχετιστική μνήμη και το  $y_k$ , είναι ένα απομνημονευμένο πρότυπο (διάνυσμα). Η συσχέτιση προτύπων που εκτελείται από το δίκτυο περιγράφεται από τη:

$$x_k \rightarrow y_k \quad k = 1, 2, \dots, q \quad (2.18)$$

όπου  $q$  είναι ο αριθμός των προτύπων που έχουν αποθηκευτεί στο δίκτυο.

### **B. Αναγνώριση Προτύπων**

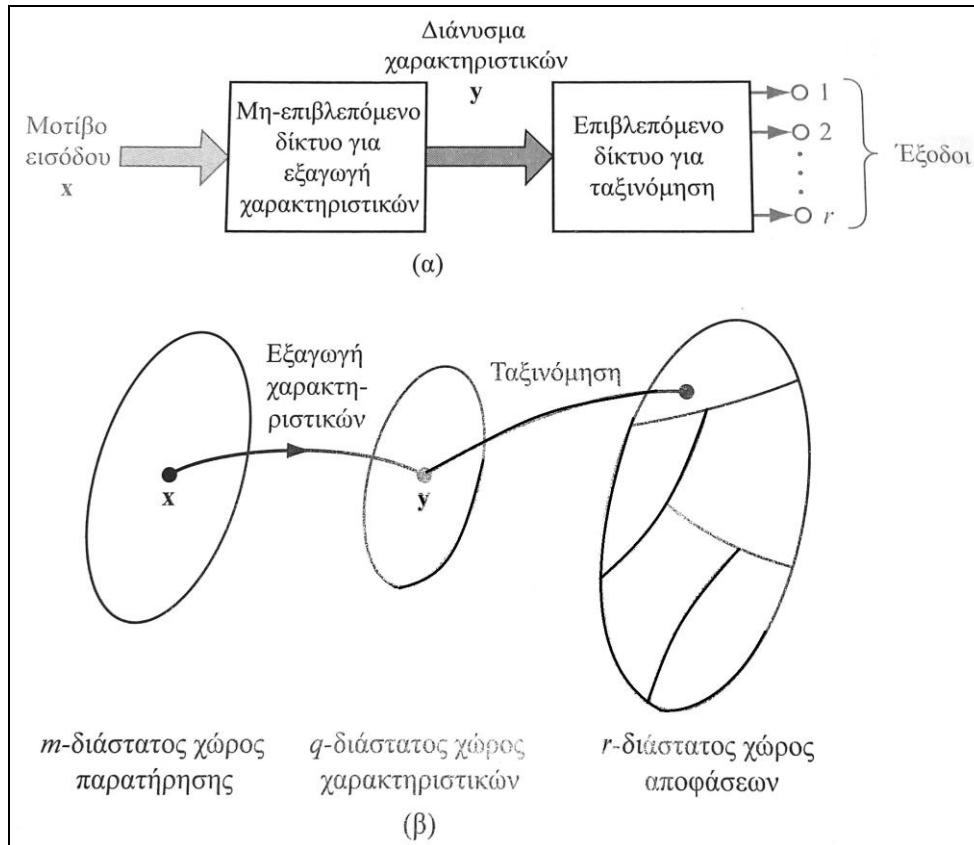
Ένα νευρωνικό δίκτυο εκτελεί αναγνώριση προτύπων αφού πρώτα περάσει από μια «σύνοδο εκπαίδευσης» κατά τη διάρκεια της οποίας παρουσιάζεται κατ' επανάληψη στο δίκτυο ένα σύνολο προτύπων εισόδου μαζί με την κατηγορία (κλάση) στην οποία ανήκει το καθένα. Στη συνέχεια, παρουσιάζεται στο δίκτυο ένα νέο πρότυπο, το οποίο δεν έχει δει ποτέ πριν, αλλά το οποίο ανήκει στον ίδιο πληθυσμό προτύπων που χρησιμοποιήθηκαν για την εκπαίδευση του δικτύου. Το δίκτυο είναι σε θέση να αναγνωρίσει την κλάση στην οποία ανήκει το συγκεκριμένο πρότυπο βάσει της πληροφορίας που έχει εξαγάγει από τα δεδομένα εκπαίδευσης [5,7].

Οι μηχανές αναγνώρισης προτύπων που χρησιμοποιούν νευρωνικά δίκτυα μπορεί να έχουν μία από τις ακόλουθες δύο μορφές:

- Η μηχανή χωρίζεται σε δύο μέρη, ένα μη επιβλεπόμενο δίκτυο για την εξαγωγή χαρακτηριστικών και ένα επιβλεπόμενο δίκτυο για την ταξινόμηση, όπως απεικονίζεται στο σύστημα του Σχ.2.15α. Η εξαγωγή χαρακτηριστικών περιγράφεται από έναν μετασχηματισμό ο οποίος αντιστοιχίζει το σημείο  $x$  σε ένα ενδιάμεσο σημείο  $y$  ενός  $q$ -διάστατου χώρου χαρακτηριστικών, όπου  $q < m$ , όπως υποδεικνύεται στο Σχ.2.15β. Η ταξινόμηση περιγράφεται ως ένας μετασχηματισμός ο οποίος

αντιστοιχίζει το ενδιάμεσο σημείο  $y$  σε μία από τις κλάσεις ενός  $r$ -διάστατου χώρου αποφάσεων, όπου  $r$  είναι ο αριθμός των διακριτών κλάσεων.

- Η μηχανή σχεδιάζεται ως ένα δίκτυο πρόσθιας τροφοδότησης χρησιμοποιώντας έναν αλγόριθμο επιβλεπόμενης μάθησης και η εργασία της εξαγωγής χαρακτηριστικών εκτελείται από τις υπολογιστικές μονάδες στα κρυφά επίπεδα του δικτύου.



Σχήμα 2.15 Η κλασική προσέγγιση της ταξινόμησης προτύπων [5].

### Γ. Προσέγγιση Συνάρτησης

Έστω μια μη γραμμική αντιστοίχιση εισόδου-εξόδου η οποία περιγράφεται από τη σχέση:

$$d = f(x) \quad (2.19)$$

όπου το διάνυσμα  $x$  είναι η είσοδος και το διάνυσμα  $d$  είναι η έξοδος. Η διανυσματική συνάρτηση  $f(\cdot)$  υποτίθεται ότι είναι άγνωστη. Σαν αντιστάθμισμα για την έλλειψη γνώσης σχετικά με τη συνάρτηση  $f(\cdot)$  μας δίνεται το ακόλουθο σύνολο χαρακτηρισμένων παραδειγμάτων:

$$J = \{(x_i, d_i)\}_{i=1}^N \quad (2.20)$$

Το ζητούμενο είναι να σχεδιάσουμε ένα νευρωνικό δίκτυο το οποίο θα προσεγγίζει την άγνωστη συνάρτηση  $f(\cdot)$  με τρόπο ώστε η συνάρτηση  $F(\cdot)$  που περιγράφει την αντιστοίχιση εισόδου-εξόδου που υλοποιείται πραγματικά από το δίκτυο, να είναι επαρκώς κοντά στην  $f(\cdot)$  για όλες τις εισόδους, όπως υποδεικνύει η σχέση:

$$\|F(x) - f(x)\| < \varepsilon \quad \text{για όλα τα } x \quad (2.21)$$

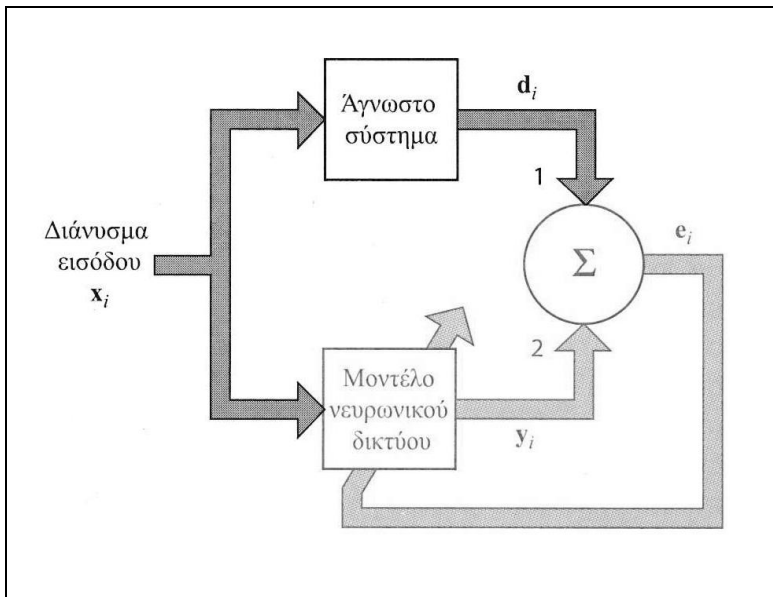
όπου  $\varepsilon$  είναι ένας μικρός θετικός αριθμός.

Εάν το μέγεθος  $N$  του δείγματος εκπαίδευσης  $G$  είναι επαρκώς μεγάλο και το δίκτυο είναι εφοδιασμένο με επαρκή αριθμό ελεύθερων παραμέτρων, το σφάλμα προσέγγισης  $\varepsilon$  μπορεί να γίνει επαρκώς μικρό για την εν λόγω εργασία.

Η δυνατότητα ενός νευρωνικού δικτύου να προσεγγίζει μια άγνωστη αντιστοίχιση εισόδου-εξόδου μπορεί να αξιοποιηθεί με δύο τρόπους:

### 1. Αναγνώριση συστημάτων

Έστω ότι η Εξ.2.19 περιγράφει τη σχέση εισόδου-εξόδου ενός άγνωστου συστήματος πολλαπλών εισόδων-πολλαπλών εξόδων (multiple-input and multiple-output - MIMO). Μπορούμε τότε να χρησιμοποιήσουμε το σύνολο χαρακτηρισμένων παραδειγμάτων της Εξ.2.20 για την εκπαίδευση ενός νευρωνικού δικτύου ως μοντέλο του συστήματος. Έστω ότι το διάνυσμα  $y_i$  συμβολίζει την πραγματική έξοδο του νευρωνικού δικτύου που παράγεται σε απόκριση προς ένα διάνυσμα εισόδου  $x_i$ . Η διαφορά μεταξύ του  $d_i$  (το οποίο σχετίζεται με το  $x_i$ ) και της εξόδου του δικτύου  $y_i$ , παρέχει το διάνυσμα σήματος σφάλματος  $e_i$ , όπως φαίνεται στο Σχ.2.16. Αυτό το σήμα σφάλματος, με τη σειρά του, χρησιμοποιείται για την προσαρμογή των ελεύθερων παραμέτρων του δικτύου, με στόχο την ελαχιστοποίηση του τετραγώνου της διαφοράς μεταξύ των εξόδων του άγνωστου συστήματος και του νευρωνικού δικτύου και υπολογίζεται για ολόκληρο το δείγμα εκπαίδευσης [5].



**Σχήμα 2.16** Σχηματικό διάγραμμα της αναγνώρισης συστημάτων. Το ΤΝΔ που εκτελεί την αναγνώριση είναι μέρος του βρόχου ανάδρασης [5].

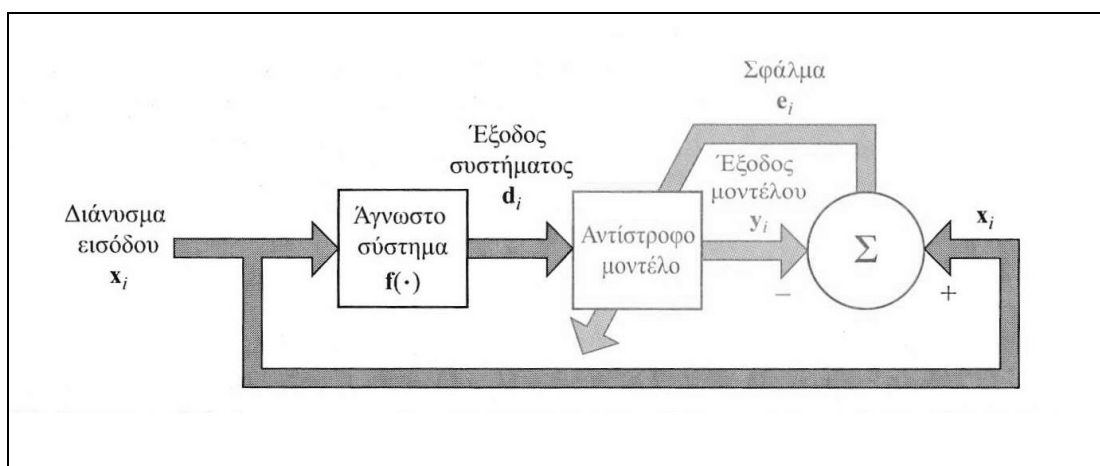
### 2. Αντίστροφη μοντελοποίηση

Υποθέστε ότι στη συνέχεια μας δίνεται ένα γνωστό άνευ μνήμης σύστημα MIMO, του οποίου η σχέση εισόδου-εξόδου περιγράφεται από την Εξ.2.19. Το ζητούμενο σε αυτή την περίπτωση είναι να κατασκευαστεί ένα αντίστροφο μοντέλο το οποίο θα παράγει το διάνυσμα  $x$  σε απόκριση προς το διάνυσμα  $d$ . Άρα, το αντίστροφο σύστημα μπορεί να περιγραφεί ως:

$$x = f^{-1}(d) \quad (2.22)$$

όπου η διανυσματική συνάρτηση  $f^{-1}(\cdot)$  συμβολίζει αναστροφή της συνάρτησης  $f(\cdot)$ . Σημειώστε, ωστόσο, ότι η  $f^{-1}(\cdot)$  δεν είναι η αντίστροφη συνάρτηση της  $f(\cdot)$ . Η χρήση του εκθέτη  $-1$  σηματοδοτεί απλώς μια αναστροφή.

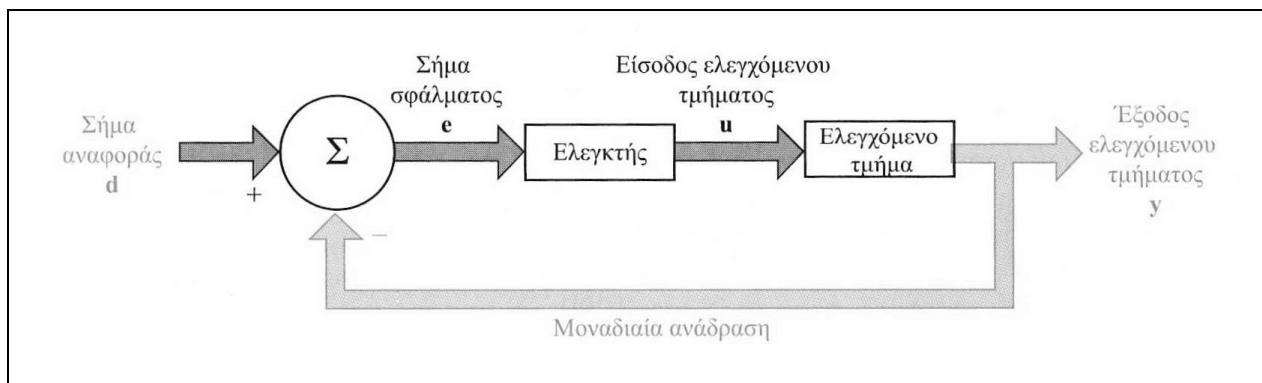
Σε πολλές περιπτώσεις που συναντώνται στην πράξη, η διανυσματική συνάρτηση  $f(\cdot)$  είναι υπερβολικά πολύπλοκη, δυσχεραίνοντας σε απαγορευτικό βαθμό τη διατύπωση της  $f^{-1}(\cdot)$ . Δοθέντος του συνόλου χαρακτηρισμένων παραδειγμάτων της Εξ.2.20, μπορούμε να υλοποιήσουμε μια προσέγγιση της  $f^{-1}(\cdot)$  μέσω του νευρωνικού δικτύου χρησιμοποιώντας το σχήμα που παρουσιάζεται στο Σχ.2.17. Στην περίπτωση που περιγράφεται εδώ, οι ρόλοι των  $x_i$  και  $d_i$ , εναλλάσσονται: Το διάνυσμα  $d_i$  χρησιμοποιείται ως είσοδος και το  $x_i$  αντιμετωπίζεται ως επιθυμητή απόκριση. Έστω ότι το διάνυσμα σήματος σφάλματος  $e_i$  συμβολίζει τη διαφορά μεταξύ του  $x_i$  και της πραγματικής εξόδου  $y_i$ , του νευρωνικού δικτύου, ή οποία παράγεται σε απόκριση προς το  $d_i$ . Όπως και στην περίπτωση της αναγνώρισης συστημάτων, αυτό το διάνυσμα σήματος σφάλματος χρησιμοποιείται για την προσαρμογή των ελεύθερων παραμέτρων του νευρωνικού δικτύου, με στόχο την ελαχιστοποίηση του τετραγώνου της διαφοράς μεταξύ των εξόδων του άγνωστου αντίστροφου συστήματος και του νευρωνικού δικτύου (υπό μια στατιστική έννοια) και υπολογίζεται για όλο το σύνολο δεδομένων εκπαίδευσης  $G$ . Συνήθως, η αντίστροφη μοντελοποίηση είναι δυσκολότερη εργασία μάθησης συγκριτικά με την αναγνώριση συστημάτων, καθώς μπορεί να μην υπάρχει μία μοναδική για αυτή [5].



**Σχήμα 2.17** Σχηματικό διάγραμμα αντίστροφης μοντελοποίησης συστημάτων. Το ΤΝΔ λειτουργώντας ως αντίστροφο μοντέλο είναι μέρος του βρόχου ανάδρασης [5].

### Α. Συστήματα Ελέγχου

Ο έλεγχος μιας μονάδας είναι άλλη μία εργασία μάθησης ιδιαίτερα κατάλληλη για υλοποίηση μέσω νευρωνικών δικτύων.



**Σχήμα 2.18** Σχηματικό διάγραμμα συστήματος ελέγχου με ανάδραση [5].

Το σύστημα που απεικονίζεται στο Σχ.2.18 χρησιμοποιεί ανάδραση γύρω από μια ελεγχόμενη μονάδα του δηλαδή, η έξοδος αυτής της μονάδας τροφοδοτείται πίσω, στην είσοδό της. Έτσι, η έξοδος  $y$  της μονάδας αφαιρείται από ένα σήμα αναφοράς  $d$  το οποίο παρέχεται από μια εξωτερική πηγή. Το παραγόμενο σήμα σφάλματος  $e$  εφαρμόζεται σε έναν νευρωνικό ελεγκτή με στόχο την προσαρμογή των ελεύθερων παραμέτρων του. Ο βασικός στόχος του ελεγκτή είναι να παρέχει κατάλληλες εισόδους στη μονάδα, έτσι ώστε η έξοδος της  $y$  να «παρακολουθεί» το σήμα αναφοράς  $d$ . Με άλλα λόγια, ο ελεγκτής πρέπει να αντιστρέφει τη συμπεριφορά εισόδου-εξόδου της μονάδας [5].

Το σήμα σφάλματος  $e$  πρέπει να περάσει από το νευρωνικό ελεγκτή πριν φτάσει στη μονάδα. Κατά συνέπεια, για να εκτελέσουμε προσαρμογές στις ελεύθερες παραμέτρους της μονάδας σύμφωνα με έναν αλγόριθμο μάθησης με διόρθωση σφάλματος, θα πρέπει να γνωρίζουμε τον πίνακα μερικών παραγώγων όπως υποδεικνύει η:

$$J = \left\{ \frac{\partial y_k}{\partial u_j} \right\}_{j,k} \quad (2.23)$$

όπου:

$y_k$  είναι ένα στοιχείο της εξόδου  $y$  της μονάδας και  
 $u_j$  είναι ένα στοιχείο της εισόδου  $u$  της μονάδας.

### **E. Πρόβλεψη Σημάτων**

Έχοντας έναν αριθμό  $N$  παρελθοντικών δειγμάτων του σήματος  $x(t-1), x(t-2), \dots, x(t-N)$  για  $t=kT, T=ct$  και  $k=1,2,\dots$  το πρόβλημα είναι να βρεθεί η τιμή  $x(t)$ . Το πρόβλημα αυτό μπορεί να λυθεί με ΤΝΔ μη επιβλεπόμενης μάθησης.



## 2.11 Το Perceptron

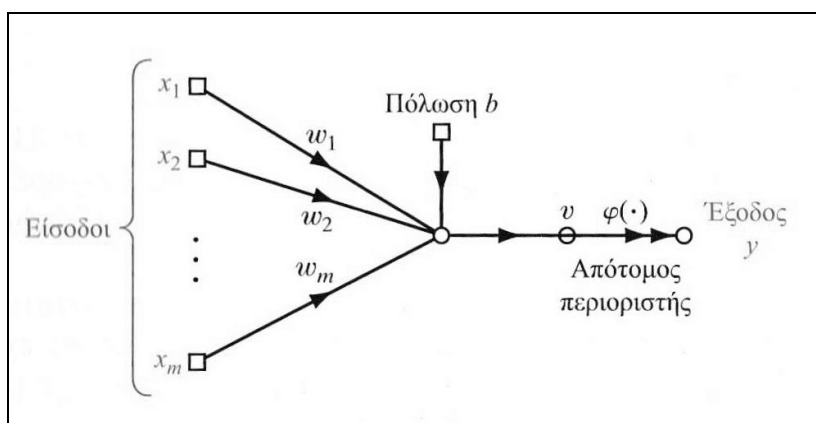
Κατά τα έτη 1943-1958, αρκετοί ερευνητές ξεχώρισαν για τις πραγματικά πρωτοποριακές συνεισφορές τους:

- McCulloch και Pitts (1943): Εισήγαγαν τη θεώρηση των νευρωνικών δικτύων ως υπολογιστικές μηχανές.
- Hebb (1949): Διατύπωσε τον πρώτο κανόνα αυτο-οργανούμενης μάθησης.
- Rosenblatt (1958): Πρότεινε το perceptron ως πρώτο μοντέλο μάθησης με τη συνδρομή ενός «εκπαιδευτή» (δηλ., επιβλεπόμενη μάθηση, ή μάθηση «με επίβλεψη»).

Το perceptron (αισθητήρας) είναι η απλούστερη δυνατή μορφή ενός νευρωνικού δικτύου που χρησιμοποιείται για την ταξινόμηση προτύπων τα οποία λέγεται ότι είναι γραμμικά διαχωρίσιμα (δηλ., πρότυπα τα οποία βρίσκονται στις αντίθετες πλευρές ενός υπερεπιπέδου). Ουσιαστικά, αποτελείται από ένα μεμονωμένο νευρώνα με προσαρμόσιμα συναπτικά βάρη και «προδιάθεση» (πόλωση, bias). Ο αλγόριθμος που χρησιμοποιείται για την προσαρμογή των ελεύθερων παραμέτρων αυτού του νευρωνικού δικτύου πρωτοεμφανίστηκε σε μια διαδικασία μάθησης που αναπτύχθηκε από τον Rosenblatt (1958, 1962) για το βασιζόμενο στο perceptron μοντέλο εγκεφάλου που πρότεινε. Πράγματι, ο Rosenblatt απέδειξε ότι εάν τα πρότυπα (διανύσματα) που χρησιμοποιούνται για την εκπαίδευση του perceptron αντλούνται από δύο γραμμικά διαχωρίσιμες κλάσεις, τότε ο αλγόριθμος του perceptron συγκλίνει και τοποθετεί τη διαχωριστική επιφάνεια απόφασης με τη μορφή ενός υπερεπιπέδου μεταξύ των δύο κλάσεων. Η απόδειξη της σύγκλισης του αλγόριθμου είναι γνωστή ως θεώρημα σύγκλισης του perceptron [5,6,8].

Το perceptron που βασίζεται σε ένα μεμονωμένο νευρώνα περιορίζεται στην ταξινόμηση προτύπων που ανήκουν σε δύο μόνο κλάσεις (υποθέσεις). Επεκτείνοντας το επίπεδο εξόδου του perceptron ώστε να περιλαμβάνει περισσότερους του ενός νευρώνες, μπορούμε να εκτελέσουμε ταξινόμηση προτύπων που ανήκουν σε περισσότερες από δύο κλάσεις. Ωστόσο, οι κλάσεις αυτές πρέπει να είναι γραμμικά διαχωρίσιμες για να δουλέψει σωστά το perceptron.

Το perceptron του Rosenblatt βασίζεται σε ένα μη γραμμικό νευρώνα συγκεκριμένα, στο μοντέλο ενός νευρώνα των McCulloch και Pitts. Ένα τέτοιο νευρωνικό μοντέλο αποτελείται από ένα γραμμικό συνδυαστή ο οποίος ακολουθείται από έναν απότομο περιοριστή, ο οποίος εκτελεί τη συνάρτηση προσημου, όπως απεικονίζεται στο Σχ.2.19. Ο κόμβος άθροισης του νευρωνικού μοντέλου υπολογίζει ένα γραμμικό συνδυασμό των εισόδων που εφαρμόζονται στις συνάψεις του, και ενσωματώνει επίσης μια εξωτερικά εφαρμόζομενη «προδιάθεση» ή «πόλωση» (bias). Το προκύπτον άθροισμα, δηλαδή, το παραγόμενο τοπικό πεδίο, εφαρμόζεται σ' έναν απότομο περιοριστή. Ως απόκριση, ο νευρώνας παράγει έξοδο ίση με +1 εάν η είσοδος του απότομου περιοριστή είναι θετική και -1 εάν είναι αρνητική [5,6,8].



Σχήμα 2.19 Γράφημα ροής σήματος του perceptron [5].

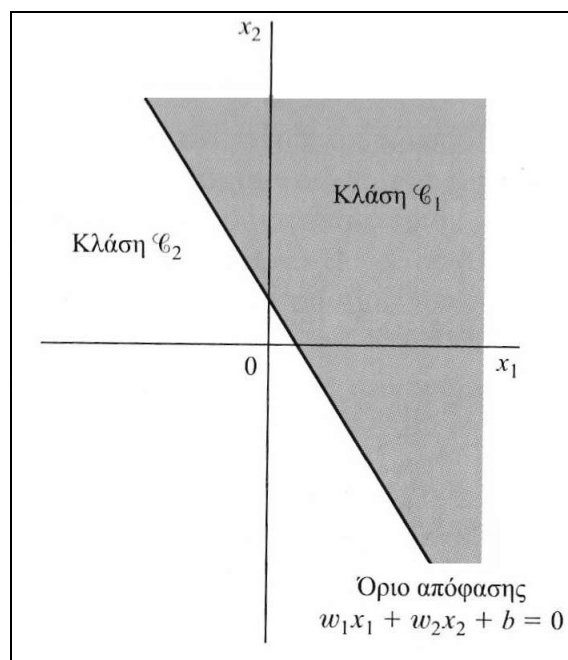
Στο γράφημα ροής σήματος του Σχ.2.19, τα συναπτικά βάρη του perceptron συμβολίζονται ως  $w_1, w_2, \dots, w_m$ . Αντίστοιχα, οι εισοδοί που εφαρμόζονται στο perceptron συμβολίζονται ως  $x_1, x_2, \dots, x_m$ . Η εξωτερικά εφαρμοζόμενη πόλωση συμβολίζεται ως  $b$ . Από το μοντέλο, βρίσκουμε ότι η είσοδος του απότομου περιοριστή, ή τοπικό πεδίο του νευρώνα, είναι:

$$v = \sum_{i=1}^m w_i x_i + b \quad (2.24)$$

Ο στόχος του perceptron είναι να ταξινομήσει σωστά το σύνολο των εξωτερικά εφαρμοζόμενων διεγέρσεων (ερεθισμάτων)  $x_1, x_2, \dots, x_m$  σε μία από δύο κλάσεις  $G_1$  ή  $G_2$ . Ο κανόνας απόφασης για την ταξινόμηση υπαγορεύει την αντιστοίχιση του σημείου που αντιπροσωπεύουν οι εισοδοί  $x_1, x_2, \dots, x_m$  στην κλάση  $G_1$  εάν η έξοδος  $y$  του perceptron είναι  $+1$ , ή στην κλάση  $G_2$  εάν είναι  $-1$ . Υπάρχουν δύο περιοχές απόφασης διαχωριζόμενες από ένα υπερεπίπεδο, το οποίο ορίζεται από την:

$$\sum_{i=1}^m w_i x_i + b = 0 \quad (2.25)$$

Αυτό παρουσιάζεται στο Σχ.2.20 για την περίπτωση δύο μεταβλητών εισόδου,  $x_1$  και  $x_2$ , για τις οποίες το όριο απόφασης παίρνει τη μορφή μιας ευθείας γραμμής. Ένα σημείο  $(x_1, x_2)$  το οποίο βρίσκεται πάνω από την οριακή γραμμή αντιστοιχίζεται στην κλάση  $G_1$ , ενώ ένα σημείο  $(x_1, x_2)$  το οποίο βρίσκεται κάτω από την οριακή γραμμή αντιστοιχίζεται στην κλάση  $G_2$ . Η επίδραση της πόλωσης  $b$  είναι απλώς να μετατοπίζει το όριο απόφασης μακριά από το σημείο αρχής των αξόνων [5,6,8].



**Σχήμα 2.20** Το υπερεπίπεδο (εδώ μια ευθεία γραμμή) ως όριο απόφασης για ένα πρόβλημα ταξινόμησης προτύπων σε δύο κλάσεις [5].

Τα συναπτικά βάρη  $w_1, w_2, \dots, w_m$  του perceptron μπορούν να προσαρμόζονται μέσω μιας επαναληπτικής διαδικασίας. Για την προσαρμογή, μπορούμε να χρησιμοποιήσουμε ένα κανόνα διόρθωσης σφαλμάτων γνωστό ως αλγόριθμο σύγκλισης του perceptron, ο οποίος ορίζεται ως εξής:

## Αλγόριθμος Σύγκλισης του perceptron

### Μεταβλητές και Παράμετροι:

- $x(n)$  = διάνυσμα εισόδων  $(m+1)$  επί  $1 = [+1, x_1(n), x_2(n), \dots, x_m(n)]^T$
- $w(n)$  = διάνυσμα βαρών  $(m+1)$  επί  $1 = [+1, w_1(n), w_2(n), \dots, w_m(n)]^T$
- $b$  = πόλωση
- $y(n)$  = πραγματική απόκριση (κβαντισμένη)
- $d(n)$  = επιθυμητή απόκριση
- $\eta$  = παράμετρος ρυθμού μάθησης, μια θετική σταθερά μικρότερη από μονάδα

1. Αρχικοποίηση: Θέσε  $w(0) = 0$ . Στη συνέχεια, εκτέλεσε τους ακόλουθους υπολογισμούς για χρονικό βήμα  $n = 1, 2$ ,
2. Ενεργοποίηση: Στο χρονικό βήμα  $n$ , ενεργοποίησε το perceptron εφαρμόζοντας το (συνεχών τιμών) διάνυσμα εισόδων  $x(n)$  και την επιθυμητή απόκριση  $d(n)$ .
3. Υπολογισμός της Πραγματικής Απόκρισης: Υπολόγισε την πραγματική απόκριση του perceptron ως:  
$$y(n) = \text{sgn}[w^T(n)x(n)]$$
  
όπου  $\text{sgn}(\cdot)$  είναι η συνάρτηση προσήμου.
4. Προσαρμογή του Διανύσματος Βαρών: Ενημέρωσε το διάνυσμα βαρών του perceptron ώστε να καταλήξεις στο:  
$$w(n+1) = w(n) + \eta[d(n) - y(n)]x(n)$$

$$\text{όπου } d(n) = \begin{cases} +1 & \text{αν το } x(n) \text{ ανήκει στην κλάση E1} \\ 0 & \text{αν το } x(n) \text{ ανήκει στην κλάση E2} \end{cases}$$

5. Συνέχιση: Αύξησε το χρονικό βήμα  $n$  κατά 1 και επέστρεψε στο βήμα 2.

Η διατύπωση του αλγόριθμου σύγκλισης του perceptron παρουσιάστηκε χωρίς αναφορά σε κάποια συνάρτηση κόστους. Η συνάρτηση κόστους είναι μια συνάρτηση η οποία επιτρέπει την εφαρμογή αναζήτησης με βάση ένα διάνυσμα κλίσεων. Ορίζουμε τη συνάρτηση κόστους του perceptron ως:

$$J(w) = \sum_{x \in X} (-w^T x) \quad (2.26)$$

όπου  $X$  είναι το σύνολο των δειγμάτων  $x$  που ταξινομούνται εσφαλμένα από ένα perceptron το οποίο χρησιμοποιεί το  $w$  ως διάνυσμα βαρών του. Εάν όλα τα δείγματα ταξινομούνται σωστά, τότε το σύνολο  $X$  είναι κενό, περίπτωση στην οποία η συνάρτηση κόστους  $J(w)$  είναι μηδέν. Επίσης, είναι διαφορίσιμη σε σχέση με το διάνυσμα βαρών  $w$  και διαφορίζοντάς τη σε σχέση με το  $w$  παίρνουμε το διάνυσμα κλίσεων:

$$\nabla J(w) = \sum_{x \in X} (-x) \quad (2.27)$$

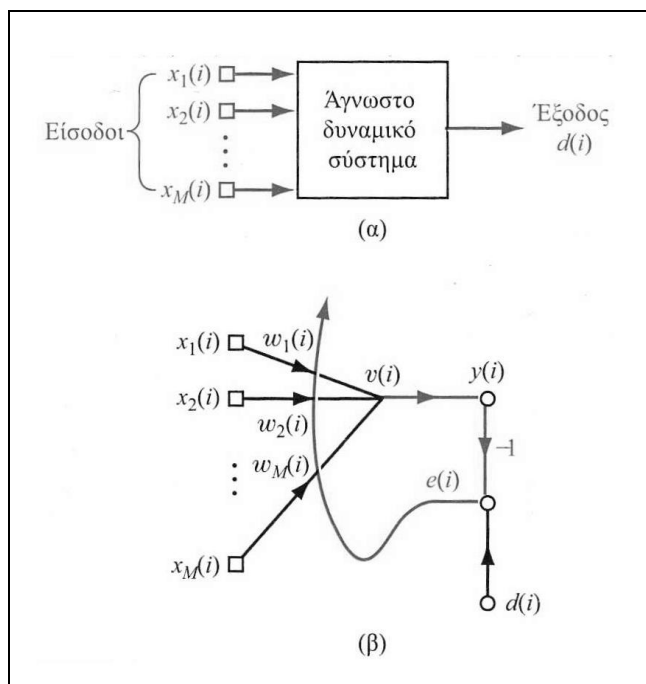
## 2.12 Ο αλγόριθμος ελαγίστων μέσων τετραγώνων για τη μέθοδο διόρθωσης σφάλματος

Στη μάθηση διόρθωσης σφάλματος όπως αναφέραμε παραπάνω, το πρόβλημα της μάθησης είναι να επιλεγούν τα συναπτικά βάρη των νευρώνων ώστε να ικανοποιείται ένα κριτήριο. Το κριτήριο αυτό είναι να ελαχιστοποιηθεί το μέσο τετραγωνικό σφάλμα μεταξύ επιθυμητών και πραγματικών αποκρίσεων [5,6,8].

Το Σχ.2.21 παρουσιάζει το γράφημα ροής σήματος για ένα προσαρμοστικό φίλτρο. Η λειτουργία του περιλαμβάνει δύο συνεχείς (στο χρόνο) διαδικασίες:

1. Μια διαδικασία φιλτραρίσματος, η οποία εκτελεί τον υπολογισμό των δύο σημάτων:
  - μια έξοδο  $y(i)$ , η οποία παράγεται σε απόκριση προς τα  $M$  στοιχεία του διανύσματος ερεθισμάτων  $x(i)$  - συγκεκριμένα,  $x_1(i), x_2(i), x_M(i)$
  - ένα σήμα σφάλματος  $e(i)$  το οποίο λαμβάνεται συγκρίνοντας την έξοδο  $y(i)$  με την αντίστοιχη έξοδο  $d(i)$  που παράγεται από το άγνωστο σύστημα. Ουσιαστικά, το  $d(i)$  δρα ως επιθυμητή απόκριση, ή σημα-στόχος.
2. Μια προσαρμοστική διαδικασία, η οποία εκτελεί την αυτόματη προσαρμογή των συναπτικών βαρών του νευρώνα σύμφωνα με το σήμα σφάλματος  $e(i)$

Η συνδυασμένη λειτουργία αυτών των δύο διαδικασιών συνιστά ένα βρόχο ανάδρασης γύρω από το νευρώνα όπως παρουσιάζεται στο Σχ.2.21β.



**Σχήμα 2.21** (α) Άγνωστο δυναμικό σύστημα. (β) Γράφημα ροής σήματος του προσαρμοστικού μοντέλου για το σύστημα. Το γράφημα ενσωματώνει ένα (μπλε) βρόχο ανάδρασης [5].

Εφόσον ο νευρώνας είναι γραμμικός η έξοδος  $y(i)$  είναι ακριβώς ίδια με το τοπικό πεδίο του νευρώνα,  $u(i)$ . Δηλαδή,

$$y(i) = v(i) = \sum_{k=1}^M w_k(i) x_k(i) \quad (2.28)$$

όπου  $w_1(i), w_2(i), \dots, w_M(i)$  είναι τα  $M$  συναπτικά βάρη του νευρώνα, μετρούμενα σε χρόνο  $i$ . Σε μορφή πίνακα, μπορούμε να εκφράσουμε την  $y(i)$  ως εσωτερικό γινόμενο των διανυσμάτων  $x(i)$  και  $w(i)$ :

$$y(i) = x^T(i)w(i) \quad (2.29)$$

όπου

$$w(i) = [w_1(i), w_2(i), \dots, w_M(i)]^T \quad (2.30)$$

Η έξοδος του νευρώνα  $y(i)$  συγκρίνεται με την αντίστοιχη έξοδο  $d(i)$  που λαμβάνεται από το άγνωστο σύστημα τη χρονική στιγμή  $i$ . Τυπικά, η  $y(i)$  είναι διαφορετική από την  $d(i)$ . Άρα, η σύγκρισή τους έχει ως αποτέλεσμα το σήμα σφάλματος:

$$e(i) = d(i) - y(i) \quad (2.31)$$

**Ο τρόπος** με τον οποίο χρησιμοποιείται το σήμα σφάλματος  $e(i)$  για τον έλεγχο των προσαρμογών στα συναπτικά βάρη του νευρώνα καθορίζεται από τη συνάρτηση κόστους που χρησιμοποιείται από τον εκάστοτε αλγόριθμο προσαρμοστικού φιλτραρίσματος [5].

#### A. ΜΕΘΟΔΟΙ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ ΧΩΡΙΣ ΠΕΡΙΟΡΙΣΜΟΥΣ

Έστω μια συνάρτηση κόστους  $E(w)$  η οποία είναι συνεχώς διαφορίσιμη συνάρτηση κάποιου άγνωστου διανύσματος βαρών (παραμέτρων)  $w$ . Η συνάρτηση  $E(w)$  αντιστοιχίζει τα στοιχεία του  $w$  σε πραγματικούς αριθμούς και αποτελεί ένα μέτρο του πώς θα πρέπει να επιλέξουμε το διάνυσμα βαρών (παραμέτρων)  $w$  για έναν αλγόριθμο προσαρμοστικού φιλτραρίσματος, έτσι ώστε να συμπεριφέρεται με βέλτιστο τρόπο [5]. Θέλουμε να βρούμε μια βέλτιστη λύση  $w^*$ , η οποία θα ικανοποιεί τη συνθήκη:

$$E(w^*) \leq E(w) \quad (2.32)$$

Ουσιαστικά πρέπει να λύσουμε ένα πρόβλημα ελαχιστοποίησης της συνάρτησης κόστους  $E(w)$  ως προς το διάνυσμα βαρών  $w$ .

Η αναγκαία συνθήκη για βελτιστότητα είναι:

$$\nabla E(w^*) = 0 \quad (2.33)$$

όπου  $\nabla$  είναι ο τελεστής κλίσης (gradient operator)

$$\nabla = \left[ \frac{\partial}{\partial w_1}, \frac{\partial}{\partial w_2}, \dots, \frac{\partial}{\partial w_M} \right]^T$$

και  $\nabla E(w)$  είναι το διάνυσμα κλίσης της συνάρτησης κόστους

$$\nabla = \left[ \frac{\partial E}{\partial w_1}, \frac{\partial E}{\partial w_2}, \dots, \frac{\partial E}{\partial w_M} \right]^T \quad (2.34)$$

**Μια κατηγορία αλγόριθμων βελτιστοποίησης χωρίς περιορισμούς** η οποία είναι ιδιαίτερα κατάλληλη για τη σχεδίαση προσαρμοστικών φίλτρων βασίζεται στο σκεπτικό της τοπικής, **επαναληπτικής κατάβασης**:

Ξεκινώντας με μια αρχική εικασία που συμβολίζεται ως  $w(0)$ , ζητείται να παραχθεί μια ακολουθία διανυσμάτων βαρών  $w(1), w(2), \dots$  τέτοια ώστε η συνάρτηση κόστους  $E(w)$  να μειώνεται σε κάθε επανάληψη του αλγόριθμου, όπως υποδεικνύει η:

$$E(w(n+1)) < E(w(n)) \quad (2.35)$$

όπου  $w(n)$  είναι η παλαιά τιμή του διανύσματος βαρών και  $w(n+1)$  είναι η ενημερωμένη τιμή του.

Υπάρχουν **τρεις μέθοδοι βελτιστοποίησης χωρίς περιορισμούς**, οι οποίες βασίζονται, με τη μια ή την άλλη μορφή, στο σκεπτικό της **επαναληπτικής κατάβασης** (Bertsekas, 1995):

1. Τη μέθοδο της πλέον απότομης κατάβασης
2. Τη μέθοδο του Newton
3. Τη μέθοδο Gauss-Newton

**Στη μέθοδο της πλέον απότομης κατάβασης (steepest descent)**, οι διαδοχικές προσαρμογές που εφαρμόζονται στο διάνυσμα βαρών  $w$  είναι προς την κατεύθυνση της πλέον απότομης κατάβασης δηλαδή, σε κατεύθυνση αντίθετη προς το διάνυσμα κλίσης  $E(w)$ :

$$g = \nabla E(w) \quad (2.36)$$

Κατά συνέπεια, ο αλγόριθμος της πλέον απότομης κατάβασης περιγράφεται από την:

$$w(n+1) = w(n) - \eta g(n) \quad (2.37)$$

όπου  $\eta$  είναι μια θετική σταθερά που αποκαλείται παράμετρος ρυθμού μάθησης και  $g(n)$  είναι το διάνυσμα κλίσης που υπολογίζεται στο σημείο  $w(n)$ . Κατά τη μετάβαση από την επανάληψη  $n$  στην  $n+1$ , ο αλγόριθμος εφαρμόζει τη διόρθωση:

$$\Delta w(n) = w(n+1) - w(n) = -\eta g(n) \quad (2.38)$$

**Η Εξίσωση (2.37-2.38) είναι στην πραγματικότητα μια διατύπωση του κανόνα διόρθωσης σφαλμάτων που περιγράψαμε παραπάνω.** Για να δείξουμε ότι ο αλγόριθμος της πλέον απότομης κατάβασης ικανοποιεί τη συνθήκη της Εξ. (2.35) για επαναληπτική κατάβαση, χρησιμοποιούμε ένα πρώτης τάξης ανάπτυγμα σειράς Taylor γύρω από την  $w(n)$  για να προσεγγίσουμε την  $E(w(n+1))$ :

$$E(w(n+1)) \approx E(w(n)) + g^T(n) \Delta w(n) \quad (2.39)$$

η χρήση του οποίου είναι επαρκής για μικρές τιμές  $\eta$ . Αντικαθιστώντας την Εξ. (2.38) σε αυτή την προσεγγιστική σχέση παίρνουμε:

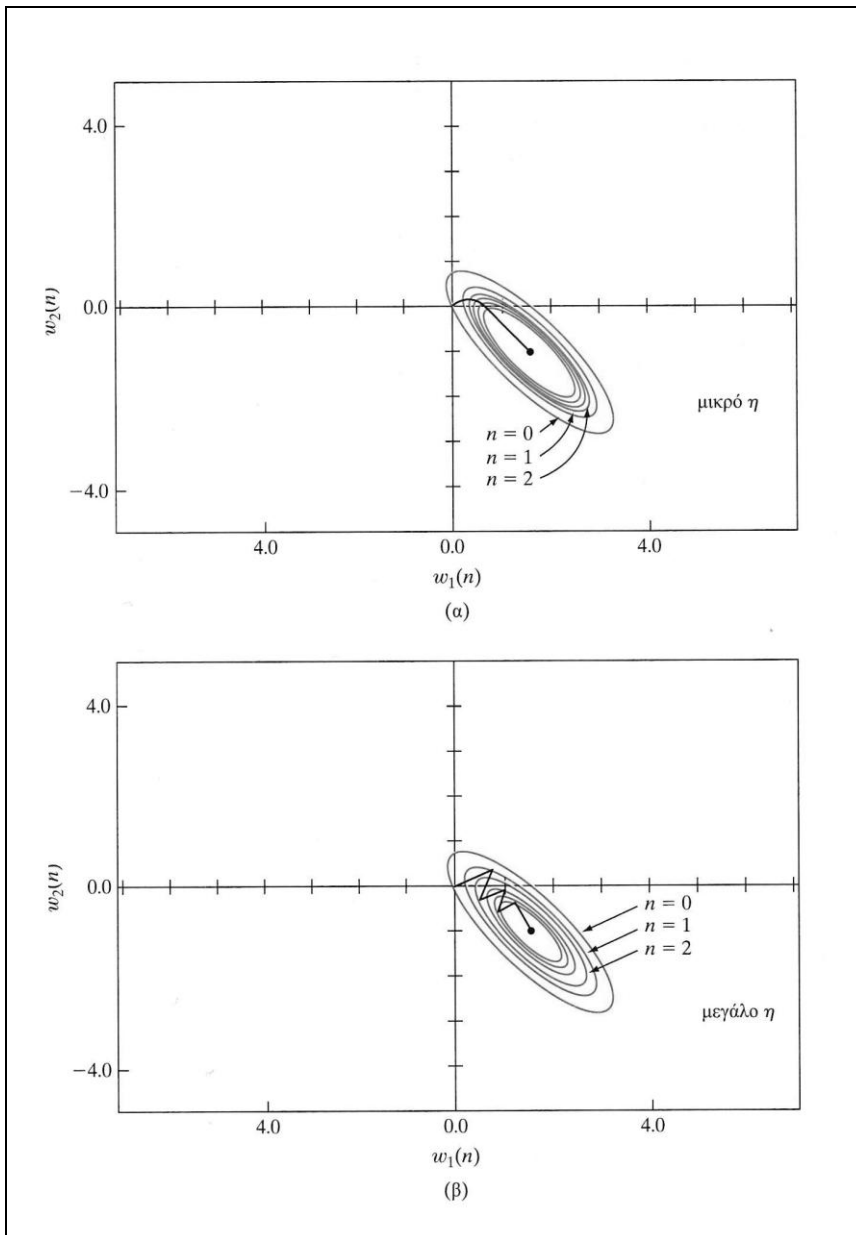
$$E(w(n+1)) \approx E(w(n)) - \eta g^T(n) g(n) = E(w(n)) - \eta \|g(n)\|^2 \quad (2.40)$$

σχέση η οποία δείχνει ότι, για θετική παράμετρο ρυθμού μάθησης  $\eta$ , η συνάρτηση κόστους μειώνεται καθώς ο αλγόριθμος προχωρά από τη μία επανάληψη στην επόμενη.

Η μέθοδος της πλέον απότομης κατάβασης συγκλίνει στη βέλτιστη λύση  $w^*$  αργά. Επιπλέον, η παράμετρος ρυθμού μάθησης  $\eta$  έχει βαθιά επίδραση στη συμπεριφορά σύγκλισης:

- Όταν το  $\eta$  είναι μικρό, η μεταβατική απόκριση του αλγόριθμου δίνει υπερβολική απόσβεση (overdamped). Η τροχιά της  $w(n)$  ακολουθεί μια ομαλή διαδρομή στο επίπεδο του  $W$ , όπως απεικονίζεται στο Σχ.2.22α.

- Όταν το  $\eta$  είναι μεγάλο, η μεταβατική απόκριση του αλγόριθμου δίνει μικρή απόσβεση (underdamped) και η τροχιά της  $w(n)$  ακολουθεί μια διαδρομή σχήματος zigzag (ταλαντωνόμενη), όπως απεικονίζεται στο Σχ.2.22β.
- Όταν το  $\eta$  υπερβαίνει μια συγκεκριμένη κρίσιμη τιμή, ο αλγόριθμος γίνεται ασταθής.



**Σχήμα 2.22** Τροχιά της μεθόδου της πλέον απότομης κατάβασης σε ένα δισδιάστατο χώρο, για δύο διαφορετικές τιμές παραμέτρου ρυθμού μάθησης  $\eta$ : (α) μικρή και (β) μεγάλη τιμή. Οι συντεταγμένες  $w_1$  και  $w_2$  είναι στοιχεία του διανύσματος βαρών  $w$ . Αμφότερα βρίσκονται στο επίπεδο του  $W$  [5].

Ως μια τεχνική βελτιστοποίησης μπορούμε να προσφύγουμε στη μέθοδο **Newton** η βασική ιδέα της οποίας είναι η ελαχιστοποίηση της τετραγωνικής προσέγγισης της συνάρτησης κόστους  $E(w)$  γύρω από το τρέχον σημείο  $w(n)$ . Αυτή η ελαχιστοποίηση εκτελείται σε κάθε επανάληψη του αλγορίθμου. Η μέθοδος Newton συγκλίνει γρήγορα ασυμπτωτικά και δεν επιδεικνύει τη συμπεριφορά zig zag που χαρακτηρίζει σε ορισμένες περιπτώσεις τη μέθοδο της πλέον απότομης κατάβασης [5].

Για να αποφύγουμε την υπολογιστική πολυπλοκότητα της μεθόδου Newton χωρίς να θέσουμε σε κίνδυνο τη συμπεριφορά της σύγκλισής της, μπορούμε **να χρησιμοποιήσουμε τη μέθοδο Gauss-Newton**.

## B. Ο ΑΛΓΟΡΙΘΜΟΣ ΕΛΑΧΙΣΤΩΝ ΜΕΣΩΝ ΤΕΤΡΑΓΩΝΩΝ (LMS)

Ο αλγόριθμος ελαχίστων μέσων τετραγώνων (least mean square, LMS), αναπτύχθηκε από τους Widrow και Hoff (1960). Η ανάπτυξη του αλγόριθμου LMS είχε πράγματι ως πηγή έμπνευσης τον αλγόριθμο του Perceptron. Αυτοί οι δύο αλγόριθμοι έχουν ένα κοινό χαρακτηριστικό: απαιτούν τη χρήση ενός γραμμικού συνδυαστή (linear combiner), γεγονός στο οποίο οφείλεται ο χαρακτηρισμός τους ως «γραμμικοί» [5,6,8].

Ο αλγόριθμος LMS κατάφερε να καθιερωθεί όχι μόνο ως η πλέον αποτελεσματική λύση για εφαρμογές προσαρμοστικού φιλτραρίσματος, αλλά επίσης και ως σημείο αναφοράς για την αξιολόγηση άλλων αλγορίθμων προσαρμοστικού φιλτραρίσματος. Οι λόγοι στους οποίους οφείλεται η εκπληκτική πορεία του είναι πολλαπλοί:

- Η υπολογιστική πολυπλοκότητα του αλγόριθμου LMS είναι γραμμική σε σχέση με τις προσαρμόσιμες παραμέτρους, πράγμα το οποίο καθιστά τον αλγόριθμο υπολογιστικώς αποτελεσματικό, ενώ ταυτόχρονα είναι επίσης αποτελεσματικός όσον αφορά την απόδοσή του.
- Είναι εύκολο να μετατραπεί σε κώδικα.
- Είναι εύρωστος αναφορικά με εξωτερικές διαταραχές.

**Ο αλγόριθμος ελαχίστων μέσων τετραγώνων (LMS) είναι σχεδιασμένος ώστε χρησιμοποιώντας μια στιγμιαία εκτίμηση για το διάνυσμα κλίσης της Εξ.2.37 της μεθόδου της πλέον απότομης κατάβασης να ελαχιστοποιεί τη στιγμιαία τιμή της συνάρτησης κόστους:**

$$E(\hat{w}) = \frac{1}{2} e^2(n) \quad (2.41)$$

όπου  $e(n)$  είναι το σήμα σφάλματος που μετρείται τη χρονική στιγμή  $n$ . Διαφορίζοντας το  $E(\hat{w})$  ως προς το διάνυσμα βαρών  $\hat{w}$  παίρνουμε:

$$\frac{\partial E(\hat{w})}{\partial \hat{w}} = e(n) \frac{\partial e(n)}{\partial w} \quad (2.42)$$

Όμοια με το φίλτρο ελαχίστων τετραγώνων, ο αλγόριθμος LMS λειτουργεί με έναν γραμμικό νευρώνα, οπότε μπορούμε να εκφράσουμε το σήμα σφάλματος ως

$$e(n) = d(n) - x^T(n) \hat{w}(n) \quad \text{άρα} \quad \frac{\partial e(n)}{\partial \hat{w}} = -x(n) \quad \text{και} \quad \frac{\partial E(\hat{w})}{\partial \hat{w}} = -x(n)e(n) \quad (2.43)$$

Χρησιμοποιώντας αυτό το τελευταίο αποτέλεσμα ως στιγμιαία εκτίμηση του διανύσματος κλίσης, μπορούμε να γράψουμε:



$$\hat{g}(n) = -x(n)e(n) \quad (2.44)$$

Οπότε, χρησιμοποιώντας την Εξ.2.44 για το διάνυσμα κλίσης της Εξ.2.37 για τη μέθοδο της πλέον απότομης κατάβασης, μπορούμε να διατυπώσουμε τον αλγόριθμο LMS ως εξής:

$$\hat{w}(n+1) = \hat{w}(n) + \eta x(n)e(n) \quad (2.45)$$

Όσο μικρότερη κάνουμε την  $\eta$ , τόσο μεγαλύτερη θα είναι η διάρκεια για την οποία ο αλγόριθμος LMS θα θυμάται παρελθοντικά δεδομένα. Κατά συνέπεια, όταν  $\eta$  είναι μικρή, ο αλγόριθμος LMS είναι ακριβής, αλλά ο ρυθμός σύγκλισής του είναι αργός.

Παρακάτω παρουσιάζεται συνοπτικά ο αλγόριθμος LMS (βάσει των Εξ. (2.43) και (2.45)). Η αρχικοποίηση του αλγόριθμου γίνεται θέτοντας απλώς την τιμή του διανύσματος βαρών  $\hat{w}(0) = 0$ .

#### Σύνοψη του Αλγόριθμου LMS

<u>Δείγμα Εκπαίδευσης</u>	Διάνυσμα σημάτων εισόδου = $x(n)$ Επιθυμητή απόκριση = $d(n)$
<u>Καθοριζόμενη από το χρήστη παράμετρος</u>	$\eta$
<u>Αρχικοποίηση</u>	Θέσε $w(0)=0$
<u>Υπολογισμός</u>	Για $n = 1, 2$ , υπολόγισε: $e(n) = d(n) - \hat{w}^T(n)x(n)$ $\hat{w}(n+1) = \hat{w}(n) + \eta x(n)e(n)$

### 2.13 Το Perceptron πολλαπλών επιπέδων

Τα βασικά χαρακτηριστικά των perceptron πολλών επιπέδων (multi layer perceptron - MLP) είναι τρία:

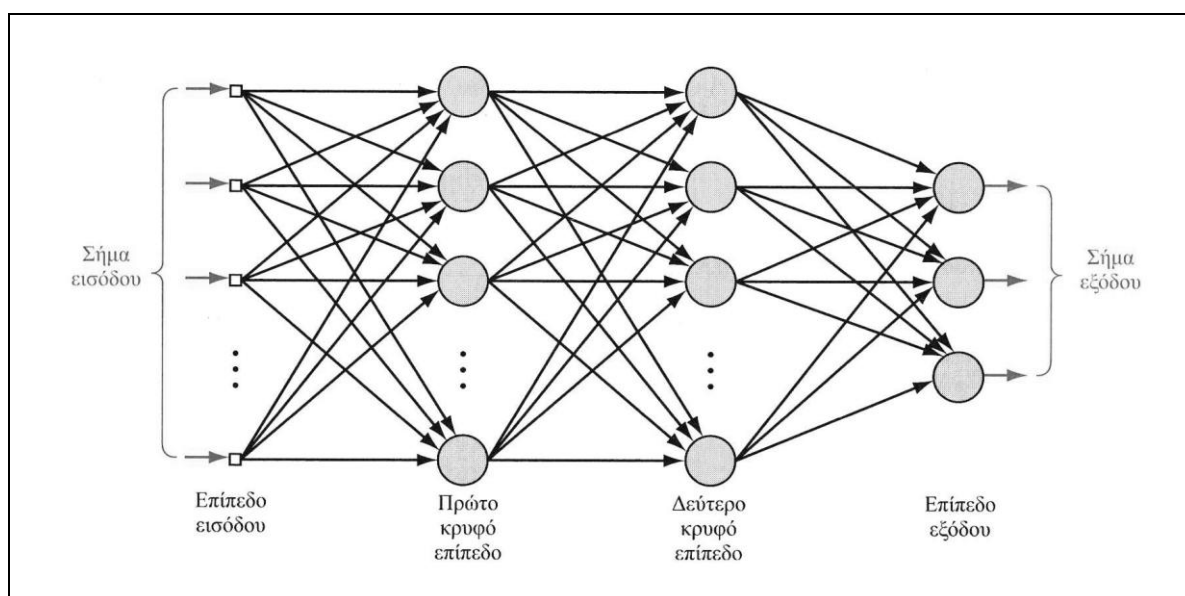
- Το μοντέλο κάθε νευρώνα στο δίκτυο περιλαμβάνει μια μη γραμμική συνάρτηση ενεργοποίησης, η οποία είναι διαφορίσιμη.
- Το δίκτυο περιέχει ένα ή περισσότερα επίπεδα τα οποία παραμένουν κρυφά για τους κόμβους των επιπέδων εισόδου και εξόδου.
- Το δίκτυο επιδεικνύει μεγάλη διασυνδεσιμότητα, ο βαθμός της οποίας καθορίζεται από τα συναπτικά βάρη του δικτύου.

Μια δημοφιλής μέθοδος για την εκπαίδευση των perceptron πολλών επιπέδων είναι ο αλγόριθμος ΒΚ, ο οποίος περιλαμβάνει τον αλγόριθμο LMS σαν ειδική περίπτωση. Η εκπαίδευση λαμβάνει χώρα σε δύο φάσεις [5]:

1. Στη φάση που εξελίσσεται προς τα εμπρός, τα συναπτικά βάρη του δικτύου είναι σταθερά και το σήμα εισόδου διαδίδεται διαμέσου του δικτύου, επίπεδο προς επίπεδο, μέχρι να φτάσει στην έξοδο. Συνεπώς, σε αυτή τη φάση, οι αλλαγές περιορίζονται στα δυναμικά ενεργοποίησης και στις εξόδους των νευρώνων του δικτύου.

2. Στη φάση που εξελίσσεται προς τα πίσω, παράγεται ένα σήμα σφάλματος μέσω της σύγκρισης της εξόδου του δικτύου με μια επιθυμητή απόκριση. Το προκύπτον σήμα σφάλματος διαδίδεται διαμέσου του δικτύου, ξανά επίπεδο προς επίπεδο, αλλά αυτή τη φορά η διάδοση γίνεται με κατεύθυνση προς τα πίσω. Σε αυτή τη δεύτερη φάση, γίνονται διαδοχικές προσαρμογές στα συναπτικά βάρη του δικτύου. Ο υπολογισμός των προσαρμογών για το επίπεδο εξόδου είναι απλή υπόθεση, ενώ για τα κρυφά επίπεδα γίνεται πολύ πιο δύσκολος.

Το Σχ. 2.23 παρουσιάζει ένα perceptron πολλών επιπέδων με δύο κρυφά επίπεδα και ένα επίπεδο εξόδου. Το συγκεκριμένο δίκτυο είναι πλήρως συνδεδεμένο. Αυτό σημαίνει ότι ένας νευρώνας σε οποιοδήποτε επίπεδο του δικτύου συνδέεται με όλους τους νευρώνες (κόμβους) του προηγούμενου επιπέδου. Η ροή σήματος διαμέσου του δικτύου προχωρά με κατεύθυνση προς τα εμπρός, από τα αριστερά προς τα δεξιά και επίπεδο προς επίπεδο.



Σχήμα 2.23 Αρχιτεκτονικός γράφος ενός perceptron πολλών επιπέδων με δύο κρυφά επίπεδα [5].

Οι νευρώνες εξόδου απαρτίζουν το επίπεδο εξόδου του δικτύου. Οι υπόλοιποι νευρώνες απαρτίζουν τα κρυφά επίπεδα του δικτύου. Έτσι, οι κρυφές μονάδες δεν αποτελούν μέρος ούτε της εξόδου, ούτε της εισόδου του δικτύου σε αυτό οφείλεται και ο χαρακτηρισμός τους ως «κρυφές». Το πρώτο κρυφό επίπεδο τροφοδοτείται από το επίπεδο εισόδου του δικτύου. Οι εξόδοι που παράγονται από το πρώτο κρυφό επίπεδο εφαρμόζονται στο επόμενο κρυφό επίπεδο η διαδικασία συνεχίζεται κατά τον ίδιο τρόπο για όλο το υπόλοιπο του δικτύου.

Κάθε κρυφός νευρώνας ή νευρώνας εξόδου ενός perceptron πολλών επιπέδων σχεδιάζεται ώστε να εκτελεί δύο υπολογισμούς:

1. Τον υπολογισμό του λειτουργικού σήματος που εμφανίζεται στην έξοδο κάθε νευρώνα, ο οποίος εκφράζεται ως μια συνεχής μη γραμμική συνάρτηση του σήματος εισόδου και των συναπτικών βαρών που σχετίζονται με αυτό το νευρώνα
2. Τον υπολογισμό μιας εκτίμησης του διανύσματος κλίσης (δηλ., των κλίσεων της επιφάνειας σφάλματος σε σχέση με τα βάρη που είναι συνδεδεμένα στις εισόδους ενός νευρώνα), ο οποίος χρειάζεται για τη φάση που εξελίσσεται προς τα πίσω στο δίκτυο.

## ΜΑΖΙΚΗ ΚΑΙ ON-LINE ΜΑΘΗΣΗ

Έστω για ένα perceptron πολλών επιπέδων, με ένα επίπεδο εισόδου (πηγαίοι κόμβοι), ένα ή περισσότερα κρυφά επίπεδα και ένα επίπεδο εξόδου αποτελούμενο από έναν ή περισσότερους νευρώνες, όπως στο Σχ.2.23 ότι το:

$$J = \{x(n), d(n)\}_{n=1}^N \quad (2.46)$$

συμβολίζει το δείγμα εκπαίδευσης που χρησιμοποιείται για την εκπαίδευση του δικτύου με επιβλεπόμενο τρόπο. Έστω ότι η  $y_j(n)$  συμβολίζει το λειτουργικό σήμα που παράγεται στην έξοδο του νευρώνα  $j$ , στο επίπεδο εξόδου, από το ερέθισμα  $x(n)$  που εφαρμόζεται στο επίπεδο εισόδου. Αντίστοιχα, το σήμα σφάλματος που παράγεται στην έξοδο του νευρώνα  $j$  ορίζεται ως:

$$e_j(n) = d_j(n) - y_j(n) \quad (2.47)$$

όπου  $d_j(n)$  είναι το  $i$ -οστό στοιχείο του διανύσματος επιθυμητών αποκρίσεων  $d(n)$ . Ακολουθώντας την ορολογία του αλγόριθμου LMS, η **στιγμιαία ενέργεια σφάλματος του νευρώνα  $j$**  ορίζεται ως:

$$E_j(n) = \frac{1}{2} e_j^2(n) \quad (2.48)$$

Αθροίζοντας τις συνεισφορές ενέργειας-σφάλματος όλων των νευρώνων στο επίπεδο εξόδου, εκφράζουμε τη **συνολική στιγμιαία ενέργεια σφάλματος ολόκληρου του δικτύου**:

$$E(n) = \sum_{j \in C} E_j(n) = \frac{1}{2} \sum_{j \in C} e_j^2(n) \quad (2.49)$$

όπου το σύνολο  $C$  περιλαμβάνει όλους τους νευρώνες στο επίπεδο εξόδου. Με το δείγμα εκπαίδευσης να αποτελείται από  $N$  παραδείγματα, η **μέση ενέργεια σφάλματος** για όλο το δείγμα εκπαίδευσης, ή αλλιώς εμπειρικό ρίσκο, ορίζεται ως:

$$E_{av}(N) = \frac{1}{N} \sum_{n=1}^N E(n) = \frac{1}{2N} \sum_{n=1}^N \sum_{j \in C} e_j^2(n) \quad (2.50)$$

Ανάλογα με το πώς εκτελείται πραγματικά η επιβλεπόμενη μάθηση στο perceptron πολλών επιπέδων, μπορούμε να αναγνωρίσουμε **δύο διαφορετικές μεθόδους** συγκεκριμένα, **τη μαζική (Batch) μάθηση** και **την on-line μάθηση** [5,6,8].

Στη **μαζική μέθοδο επιβλεπόμενης μάθησης**, οι προσαρμογές στα συναπτικά βάρη του perceptron πολλών επιπέδων εκτελούνται μετά από την παρουσίαση του συνόλου των  $N$  παραδειγμάτων του δείγματος εκπαίδευσης  $J$ , τα οποία συνιστούν μια εποχή εκπαίδευσης.

Με άλλα λόγια, η συνάρτηση κόστους για τη μαζική μάθηση ορίζεται από τη μέση ενέργεια σφάλματος,  $E_{av}$ . Οι προσαρμογές στα συναπτικά βάρη του perceptron πολλών επιπέδων γίνονται σε βάση **εποχή προς εποχή**. Κατά συνέπεια, μπορούμε να παράγουμε μια υλοποίηση της καμπύλης μάθησης απεικονίζοντας το  $E_{av}$  έναντι του αριθμού των εποχών, όπου, για κάθε εποχή εκπαίδευσης, τα παραδείγματα στο δείγμα εκπαίδευσης  $J$  είναι με τυχαία σειρά (ανακατεμένα). Η καμπύλη μάθησης παράγεται υπολογίζοντας στατιστικούς μέσους συνόλου για επαρκώς μεγάλο αριθμό τέτοιων υλοποιήσεων, όπου κάθε υλοποίηση εκτελείται για ένα *διαφορετικό σύνολο αρχικών, τυχαία επιλεγμένων συνθηκών*. Η μαζική μάθηση μπορεί να αντιμετωπιστεί ως μια μορφή στατιστικού συμπερασμού. Για το λόγο αυτό είναι ιδιαίτερα κατάλληλη για την επίλυση προβλημάτων μη γραμμικής παλινδρόμησης [5,6,8].

Με την **online μέθοδο επιβλεπόμενης μάθησης**, οι προσαρμογές στα συναπτικά βάρη του perceptron πολλών επιπέδων εκτελούνται σε βάση **παράδειγμα προς παράδειγμα**. Η συνάρτηση κόστους προς ελαχιστοποίηση είναι, συνεπώς, η συνολική στιγμιαία ενέργεια σφάλματος  $E(n)$ .

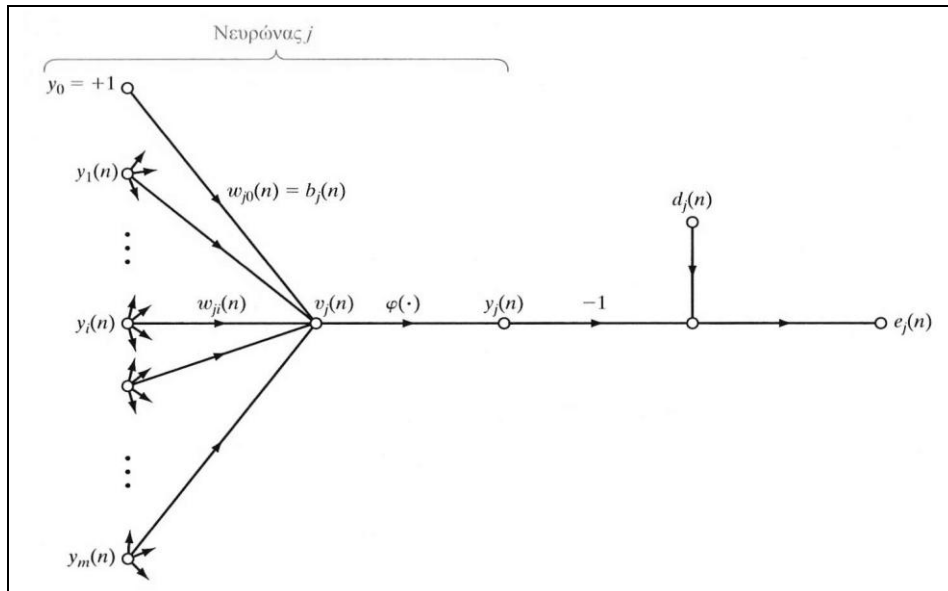
Για ένα δεδομένο σύνολο αρχικών συνθηκών, μπορούμε να λάβουμε μια μεμονωμένη υλοποίηση της καμπύλης μάθησης απεικονίζοντας την τελική τιμή  $E(N)$  σε συνάρτηση με τον αριθμό των εποχών που χρησιμοποιούνται κατά τη διάρκεια της συνόδου εκπαίδευσης, όπου, όπως και πριν, τα παραδείγματα εκπαίδευσης αναδιοργανώνονται με τυχαία σειρά μετά από κάθε εποχή. Όπως και με τη μαζική μάθηση, η καμπύλη μάθησης για την online μάθηση λαμβάνεται υπολογίζοντας στατιστικούς μέσους συνόλου τέτοιων υλοποιήσεων για ένα επαρκώς μεγάλο αριθμό αρχικών συνθηκών, τυχαία επιλεγμένων. Φυσικά, για μια δεδομένη δομή δικτύου, η καμπύλη μάθησης που επιτυγχάνεται με την online μάθηση θα είναι αρκετά διαφορετική από αυτή για τη μαζική μάθηση [5,6,8].

Η online μάθηση είναι εξαιρετικά δημοφιλής για την επίλυση προβλημάτων ταξινόμησης προτύπων, για δύο σημαντικούς πρακτικούς λόγους:

- Η online μάθηση είναι απλή στην υλοποίηση.
- Παρέχει αποτελεσματικές λύσεις σε μεγάλης κλίμακας και αυξημένης δυσκολίας προβλήματα ταξινόμησης προτύπων.

## Ο ΑΛΓΟΡΙΘΜΟΣ BACK PROPAGATION (BK)

Για να περιγράψουμε αυτό τον αλγόριθμο, θα ξεκινήσουμε από το Σχ.2.24 το οποίο απεικονίζει το νευρώνα να τροφοδοτείται από ένα σύνολο λειτουργικών σημάτων που παράγονται από ένα επίπεδο νευρώνων στα αριστερά του.



**Σχήμα 2.24** Γράφημα ροής σήματος που περιγράφει τις λεπτομέρειες του νευρώνα εξόδου  $j$  [5].

Το τοπικό πεδίο  $v_j(n)$  που παράγεται στην είσοδο της συνάρτησης ενεργοποίησης που σχετίζεται με το νευρώνα  $j$  είναι:

$$v_j(n) = \sum_{i=0}^m w_{ji}(n) y_i(n) \quad (2.51)$$

όπου  $m$  είναι ο συνολικός αριθμός εισόδων (εξαιρουμένης της πόλωσης) που εφαρμόζονται στο νευρώνα  $j$ . Το συναπτικό βάρος  $w_{j0}$  (αντιστοιχεί στη σταθερή είσοδο  $y_0 = +1$ ) ισούται με την πόλωση  $b_j$ , που εφαρμόζεται στο νευρώνα  $j$ . Άρα, το λειτουργικό σήμα  $y_j(n)$  που εμφανίζεται στην έξοδο του νευρώνα  $j$  κατά την επανάληψη  $n$  είναι:

$$y_j(n) = \varphi_j(v_j(n)) \quad (2.52)$$

Όμοια με τον αλγόριθμο LMS που εξετάστηκε προηγουμένως, ο αλγόριθμος BK εφαρμόζει μια διόρθωση  $\Delta w_{ji}(n)$  στο συναπτικό βάρος  $w_{ji}(n)$ , η οποία είναι ανάλογη με τη μερική παράγωγο  $\partial E(n) / \partial w_{ji}(n)$ . Σύμφωνα με τον κανόνα αλυσίδας του διαφορικού λογισμού, μπορούμε να εκφράσουμε αυτή την κλίση ως:

$$\frac{\partial E(n)}{\partial w_{ji}(n)} = \frac{\partial E(n)}{\partial e_j(n)} \frac{\partial e_j(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)} \frac{\partial v_j(n)}{\partial w_{ji}(n)} \quad (2.53)$$

Η μερική παράγωγος  $\partial E(n) / \partial w_{ji}(n)$  αντιπροσωπεύει ένα συντελεστή ευαισθησίας, ο οποίος καθορίζει την κατεύθυνση της αναζήτησης στο χώρο βαρών για το συναπτικό βάρος  $w_{ji}$ .

Διαφορίζοντας και τις δύο πλευρές της Εξ.2.49 ως προς  $e_j(n)$ , λαμβάνουμε:

$$\frac{\partial E(n)}{\partial e_j(n)} = e_j(n) \quad (2.54)$$

Διαφορίζοντας και τις δύο πλευρές της Εξ.2.47 ως προς  $y_j(n)$ , λαμβάνουμε:

$$\frac{\partial e_j(n)}{\partial y_j(n)} = -1 \quad (2.55)$$

Στη συνέχεια, διαφορίζοντας την Εξ.2.52 ως προς  $v_j(n)$ , λαμβάνουμε:

$$\frac{\partial y_j(n)}{\partial v_j(n)} = \varphi'(v_j(n)) \quad (2.56)$$

όπου η χρήση πρώτης παραγώγου (στη δεξιά πλευρά) σηματοδοτεί διαφορίση ως προς το όρισμα. Τέλος, διαφορίζοντας την Εξ.2.51 ως προς  $w_{ji}(n)$  παίρνουμε:

$$\frac{\partial v_j(n)}{\partial w_{ji}(n)} = y_i(n) \quad (2.57)$$

Η χρήση των Εξ.2.54 και 2.57 στην Εξ.2.53 δίνει:

$$\frac{\partial E(n)}{\partial w_{ji}(n)} = -e_j(n)\varphi'(v_j(n))y_i(n) \quad (2.58)$$

Η διόρθωση  $\Delta w_{ji}(n)$  που εφαρμόζεται στο  $w_{ji}(n)$  ορίζεται από **τον κανόνα Δέλτα**:

$$\Delta w_{ji}(n) = -\eta \frac{\partial E(n)}{\partial w_{ji}(n)} \quad (2.59)$$

όπου  $\eta$  είναι η παράμετρος ρυθμού μάθησης του αλγόριθμου ΒΚ. Η χρήση της Εξ. (2.58) στην Εξ. (2.59) δίνει:

$$\Delta w_{ji}(n) = \eta \delta_j(n) y_i(n) \quad (2.60)$$

**όπου η τοπική κλίση  $\delta_j(n)$  ορίζεται από την**

$$\delta_j(n) = \frac{\partial E(n)}{\partial v_j(n)} = \frac{\partial E(n)}{\partial e_j(n)} \frac{\partial e_j(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)} = e_j(n)\varphi'(v_j(n)) \quad (2.61)$$

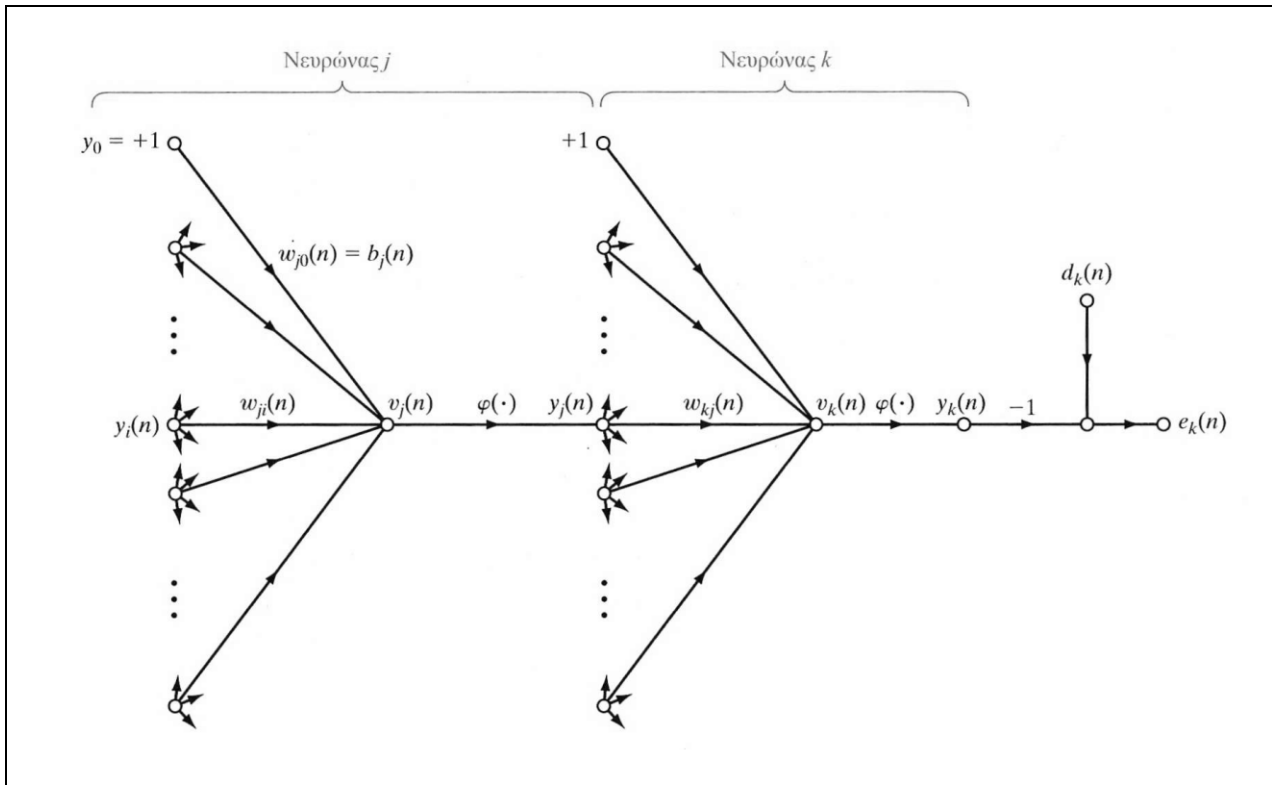
Μπορούμε να αναγνωρίζουμε δύο ξεχωριστές περιπτώσεις, ανάλογα με το πού βρίσκεται ο νευρώνας  $j$  στο δίκτυο [5,6,8].

Περίπτωση 1: Ο Νευρώνας  $j$  είναι ένας κόμβος εξόδου

Όταν ο νευρώνας  $j$  βρίσκεται στο επίπεδο εξόδου του δικτύου, τροφοδοτείται με τη δική του επιθυμητή απόκριση. Μπορούμε να χρησιμοποιήσουμε την Εξ. (2.47) για να υπολογίσουμε το σήμα σφάλματος  $e_j(n)$  που σχετίζεται με αυτό το νευρώνα. Έχοντας καθορίσει το  $e_j(n)$  είναι εύκολο να υπολογίσουμε την τοπική κλίση  $\delta_j(n)$  χρησιμοποιώντας την Εξ (2.61).

Περίπτωση 2: Ο Νευρώνας j είναι ένας κρυφός κόμβος

Όταν ο νευρώνας j βρίσκεται σε ένα κρυφό επίπεδο του δικτύου, δεν υπάρχει καθορισμένη επιθυμητή απόκριση για αυτόν. Οπότε, το σήμα σφάλματος για ένα κρυφό νευρώνα θα πρέπει να καθοριστεί αναδρομικά, δουλεύοντας προς τα πίσω, βάσει των σημάτων σφάλματος όλων των νευρώνων με τους οποίους συνδέεται άμεσα αυτός ο κρυφός νευρώνας. Το Σχ.2.25 απεικονίζει το νευρώνα j ως ένα κρυφό κόμβο του δικτύου.



**Σχήμα 2.25** Γράφημα ροής σήματος που περιγράφει τις λεπτομέρειες του νευρώνα εξόδου k ο οποίος συνδέεται στον κρυφό νευρώνα j [5].

Σύμφωνα με την Εξ. (2.61), μπορούμε να επαναπροσδιορίσουμε την τοπική κλίση  $\delta_j(n)$  για τον κρυφό νευρώνα j ως:

$$\begin{aligned} \delta_j(n) &= -\frac{\partial E(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)} = \\ &= -\frac{\partial E(n)}{\partial y_j(n)} \phi'_j(v_j(n)) \end{aligned} \quad \text{(ο νευρώνας j είναι κρυφός)} \quad (2.62)$$

όπου στη δεύτερη γραμμή έχουμε χρησιμοποιήσει την Εξ. (2.56). Για να υπολογίσουμε τη μερική παράγωγο  $\partial E(n) / \partial y_j(n)$ , μπορούμε να προχωρήσουμε ως εξής: Από το Σχ.2.25, βλέπουμε ότι:

$$\partial E(n) = \frac{1}{2} \sum_{k \in C} e_k^2(n) \quad \text{(ο νευρώνας k είναι ένας κόμβος εξόδου)} \quad (2.63)$$

η οποία είναι η Εξ. (2.49) με το δείκτη k στη θέση του δείκτη j. Διαφορίζοντας την Εξ. (2.63) ως προς το λειτουργικό σήμα  $y_j(n)$ , λαμβάνουμε:

$$\frac{\partial E(n)}{\partial y_j(n)} = \sum_k e_k \frac{\partial e_k(n)}{\partial y_j(n)} \quad (2.64)$$

Στη συνέχεια, χρησιμοποιούμε τον κανόνα αλυσίδας για τη μερική παράγωγο  $\partial e_k(n) / \partial y_j(n)$  και αναδιατυπώνουμε την Εξ. (2.64) στην ισοδύναμη μορφή:

$$\frac{\partial E(n)}{\partial y_j(n)} = \sum_k e_k(n) \frac{\partial e_k(n)}{\partial v_k(n)} \frac{\partial v_k(n)}{\partial y_j(n)} \quad (2.65)$$

Επίσης, από το Σχ.2.25, παρατηρούμε ότι:

$$e_k(n) = d_k(n) - y_j(n) = d_k(n) - \varphi_k(v_k(n)) \quad (\text{ο νευρώνας } k \text{ είναι ένας κόμβος εξόδου}) \quad (2.66)$$

Άρα,

$$\frac{\partial e_k(n)}{\partial v_k(n)} = -\varphi'_k(v_k(n)) \quad (2.67)$$

Από το Σχ.2.25 παρατηρούμε επίσης ότι για το νευρώνα  $k$ , το τοπικό πεδίο είναι:

$$v_k(n) = \sum_{j=0}^m w_{kj}(n) y_j(n) \quad (2.68)$$

όπου  $m$  είναι ο συνολικός αριθμός εισόδων (εξαιρουμένης της πόλωσης) που εφαρμόζονται στο νευρώνα  $k$ . Το συναπτικό βάρος  $w_{k0}(n)$  ισούται με την πόλωση  $b_k(n)$  που εφαρμόζεται στο νευρώνα  $k$  και η αντίστοιχη είσοδος είναι σταθερή σε τιμή +1. Διαφορίζοντας την Εξ. (2.68) ως προς  $y_j(n)$  παίρνουμε:

$$\frac{\partial v_k(n)}{\partial y_j(n)} = w_{kj}(n) \quad (2.69)$$

Χρησιμοποιώντας τις Εξ. (2.67) και (2.69) στην Εξ. (2.65), παίρνουμε την επιθυμητή μερική παράγωγο

$$\begin{aligned} \frac{\partial E(n)}{\partial y_j(n)} &= -\sum_k e_k(n) \varphi'_k(v_k(n)) w_{kj}(n) = \\ &= -\sum_k \delta_k(n) w_{kj}(n) \end{aligned} \quad (2.70)$$

όπου, στη δεύτερη γραμμή έχουμε χρησιμοποιήσει τον ορισμό της τοπικής κλίσης  $\delta_k(n)$  που δόθηκε στην Εξ. (2.61), με το δείκτη  $k$  στη θέση του  $j$ .

Χρησιμοποιώντας την Εξ. (2.70) στην Εξ. (2.62), παίρνουμε τον τύπο οπισθοδιάδοσης για την τοπική κλίση  $\delta_j(n)$ , ο οποίος περιγράφεται από την

$$\delta_j(n) = \varphi'_j(v_j(n)) \sum_k \delta_k(n) w_{kj}(n) \quad (\text{ο νευρώνας } j \text{ είναι κρυφός}) \quad (2.71)$$



Η διόρθωση  $\Delta w_{kj}(n)$  που εφαρμόζεται στο συναπτικό βάρος που συνδέει το νευρώνα  $i$  με το νευρώνα  $j$  ορίζεται από τον **κανόνα Δέλτα**:

$$\begin{pmatrix} \text{Διόρθωση} \\ \text{βάρους} \\ \Delta w_{ji}(n) \end{pmatrix} = \begin{pmatrix} \text{παράμετρος} \\ \text{ρυθμού μάθησης} \\ \eta \end{pmatrix} \times \begin{pmatrix} \text{τοπική} \\ \text{κλίση} \\ \delta_j(n) \end{pmatrix} \times \begin{pmatrix} \text{σήμα εισόδου} \\ \text{του νευρώνα } j \\ y_i(n) \end{pmatrix} \quad (2.72)$$

Επίσης, η τοπική κλίση  $\delta_j(n)$  εξαρτάται από το εάν ο νευρώνας  $j$  είναι ένας κόμβος εξόδου ή ένας κρυφός κόμβος:

1. Εάν ο νευρώνας  $j$  είναι ένας κόμβος εξόδου, το  $\delta_j(n)$  ισούται με το γινόμενο της παραγώγου  $\varphi'_j(v_j(n))$  και του σήματος σφάλματος  $e_j(n)$ , αμφότερα εκ των οποίων συσχετίζονται με το νευρώνα  $j$  (δες την Εξ. (2.61)).
2. Εάν ο νευρώνας  $j$  είναι ένας κρυφός κόμβος, το  $\delta_j(n)$  ισούται με το γινόμενο της σχετιζόμενης παραγώγου  $\varphi'_j(v_j(n))$  και του σταθμισμένου αθροίσματος των  $\delta$  που υπολογίζονται για τους νευρώνες του επόμενου κρυφού επιπέδου ή επιπέδου εξόδου και οι οποίοι συνδέονται με το νευρώνα  $j$  (δες την Εξ. (2.71)).

Κατά την εφαρμογή του αλγόριθμου BK, έχουμε δύο διαφορετικά περάσματα υπολογισμών. Το πρώτο πέρασμα αναφέρεται ως πέρασμα με κατεύθυνση προς τα εμπρός και το δεύτερο αναφέρεται ως πέρασμα με κατεύθυνση προς τα πίσω.

Στο πέρασμα με κατεύθυνση προς τα εμπρός, τα συναπτικά βάρη παραμένουν αμετάβλητα σε όλη την έκταση του δικτύου και τα λειτουργικά σήματα του δικτύου υπολογίζονται σε βάση νευρώνα προς νευρώνα. Το λειτουργικό σήμα που εμφανίζεται στην έξοδο του νευρώνα  $j$  υπολογίζεται ως:

$$y_j(n) = \varphi(v_j(n)) \quad (2.73)$$

όπου  $v_j(n)$  είναι το τοπικό πεδίο του νευρώνα  $j$ , το οποίο ορίζεται ως

$$v_j(n) = \sum_{i=0}^m w_{ji}(n) y_i(n) \quad (2.74)$$

όπου  $m$  είναι ο συνολικός αριθμός εισόδων (εξαιρουμένης της πόλωσης) που εφαρμόζονται στο νευρώνα  $j$ . Το  $w_{ji}(n)$  είναι το συναπτικό βάρος που συνδέει το νευρώνα  $i$  με το νευρώνα  $j$  και  $y_i(n)$  είναι ένα σήμα εισόδου για το νευρώνα  $j$ , ή, ισοδύναμα, το λειτουργικό σήμα που εμφανίζεται στην έξοδο του νευρώνα  $i$ . Εάν ο νευρώνας  $j$  είναι στο πρώτο κρυφό επίπεδο του δικτύου, τότε  $m=m_0$  και ο δείκτης  $i$  αναφέρεται στον  $i$ -οστό τερματικό κόμβο εισόδου του δικτύου, για τον οποίο γράφουμε:

$$y_i(n) = x_i(n) \quad (2.75)$$

όπου  $x_i(n)$  είναι το  $i$ -οστό στοιχείο του διανύσματος εισόδου (πρότυπο). Εάν, από την άλλη, ο νευρώνας  $j$  είναι στο επίπεδο εξόδου του δικτύου, τότε  $m=m_L$  και ο δείκτης  $j$  αναφέρεται στον  $j$ -οστό τερματικό κόμβο εξόδου του δικτύου, για τον οποίο γράφουμε:

$$y_j(n) = o_j(n) \quad (2.76)$$

όπου  $o_j(n)$  είναι το  $j$ -οστό στοιχείο του διανύσματος εξόδου του perceptron πολλών επιπέδων. Αυτή η έξοδος συγκρίνεται με την επιθυμητή απόκριση  $d_j(n)$  και λαμβάνεται το σήμα σφάλματος  $e_j(n)$  για τον  $j$ -οστό νευρώνα εξόδου. Έτσι, η φάση του υπολογισμού με κατεύθυνση προς τα εμπρός ξεκινά στο πρώτο κρυφό επίπεδο, παρουσιάζοντάς του το διάνυσμα εισόδων και τερματίζει στο επίπεδο εξόδου, υπολογίζοντας το σήμα σφάλματος για κάθε νευρώνα αυτού του επιπέδου [5,6,8].

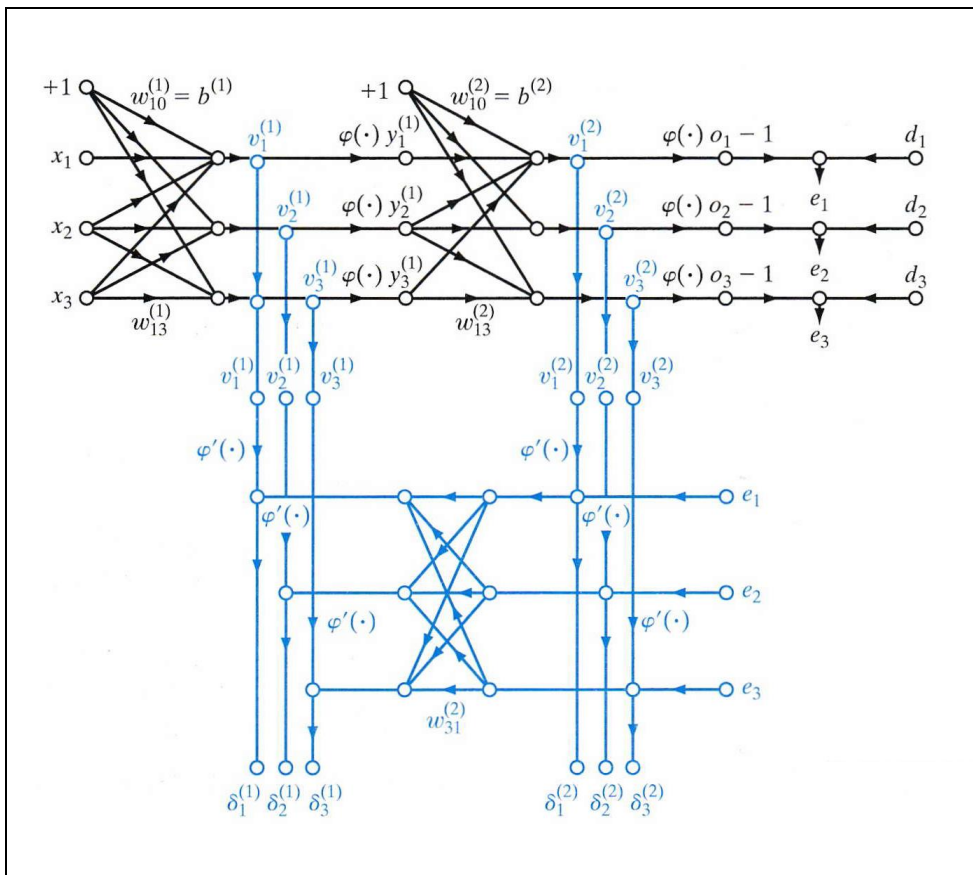
Το πέρασμα με κατεύθυνση προς τα πίσω ξεκινά στο επίπεδο εξόδου, στέλνοντας τα σήματα σφάλματος προς τα αριστερά, σε όλα τα επίπεδα του δικτύου, επίπεδο προς επίπεδο, και υπολογίζοντας αναδρομικά το  $\delta$  (δηλ., την τοπική κλίση) για κάθε νευρώνα. Αυτή η αναδρομική διαδικασία επιτρέπει στα συναπτικά βάρη του δικτύου να υφίστανται μεταβολές σύμφωνα με τον κανόνα Δέλτα της Εξ. (2.72). Για ένα νευρώνα ευρισκόμενο στο επίπεδο εξόδου, το  $\delta$  είναι απλώς ίσο με το σήμα σφάλματος αυτού του νευρώνα πολλαπλασιαζόμενο επί την πρώτη παράγωγο της μη γραμμικότητάς του. Συνεπώς, χρησιμοποιούμε την Εξ.(2.72) για να υπολογίσουμε τις μεταβολές στα βάρη όλων των συνδέσεων που τροφοδοτούν το επίπεδο εξόδου. Με δεδομένα τα  $\delta$  για τους νευρώνες του επιπέδου εξόδου, στη συνέχεια χρησιμοποιούμε με την Εξ.(2.71) για να υπολογίσουμε τα  $\delta$  για όλους τους νευρώνες στο προτελευταίο επίπεδο και, κατά συνέπεια, τις μεταβολές στα βάρη όλων των συνδέσεων που το τροφοδοτούν. Αυτός ο αναδρομικός υπολογισμός συνεχίζεται, επίπεδο προς επίπεδο, διαδίδοντας τις μεταβολές όλων των συναπτικών βαρών στο δίκτυο [5,8].

Μια απλή μέθοδος αύξησης του ρυθμού μάθησης με ταυτόχρονη αποφυγή του κινδύνου της αστάθειας είναι η τροποποίηση του κανόνα Δέλτα της Εξ. (4.15) συμπεριλαμβάνοντας έναν όρο ορμής, όπως υποδεικνύει η

$$\Delta w_{ji}(n) = a\Delta w_{ji}(n-1) + \eta\delta_j(n)y_i(n) \quad (2.78)$$

όπου  $a$  είναι συνήθως ένας θετικός αριθμός και αποκαλείται σταθερά ορμής. **Η παραπάνω εξίσωση αποκαλείται γενικευμένος κανόνας Δέλτα.** Περιλαμβάνει τον κανόνα Δέλτα της Εξ. (2.59-2.60) ως ειδική περίπτωση (δηλ.,  $a = 0$ ).

Το Σχ.2.23 παρουσιάζει την αρχιτεκτονική διάταξη ενός perceptron πολλών επιπέδων. Το αντίστοιχο γράφημα ροής σήματος για τη μάθηση με οπισθοδιάδοση σφάλματος, το οποίο ενσωματώνει και τις δύο (προς τα εμπρός και προς τα πίσω) φάσεις των υπολογισμών που εμπλέκονται στη διαδικασία μάθησης, παρουσιάζεται στο Σχ. 2.26 για την περίπτωση όπου  $L = 2$  και  $m_0 = m_1 = m_2 = 3$ . Το πάνω μέρος του γραφήματος ροής σήματος αφορά το πέρασμα με κατεύθυνση προς τα εμπρός. Το κάτω μέρος του γραφήματος ροής σήματος αφορά το πέρασμα με κατεύθυνση προς τα πίσω.

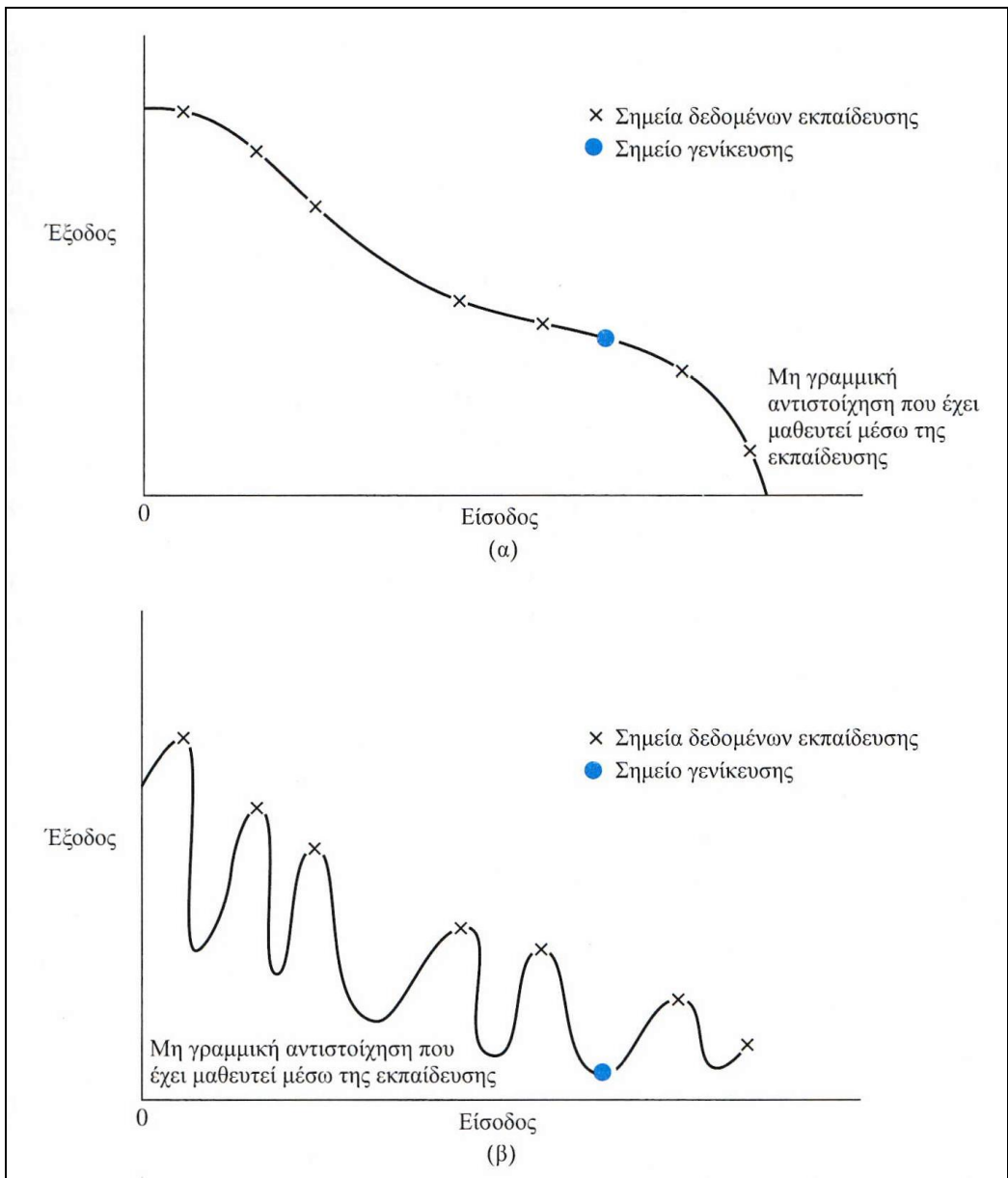


**Σχήμα 2.26** Συνοπτική γραφική απεικόνιση της διαδικασίας μάθησης με οπισθοδιάδοση σφάλματος. Επάνω μέρος του γραφήματος: Το πέρασμα με κατεύθυνση προς τα εμπρός. Κάτω μέρος του γραφήματος: Το πέρασμα με κατεύθυνση προς τα πίσω [5].

## ΓΕΝΙΚΕΥΣΗ

Ένα δίκτυο λέγεται ότι γενικεύει καλά όταν η αντιστοίχιση εισόδου-εξόδου που υπολογίζεται από το δίκτυο είναι σωστή για δεδομένα ελέγχου που δεν χρησιμοποιήθηκαν ποτέ κατά τη δημιουργία ή την εκπαίδευση του δικτύου. Το Σχ.2.27 απεικονίζει πώς μπορεί να λάβει χώρα η γενίκευση σε ένα υποθετικό δίκτυο. Η μη γραμμική αντιστοίχιση εισόδου-εξόδου που αναπαριστά η καμπύλη του σχήματος υπολογίζεται από το δίκτυο σαν αποτέλεσμα της μάθησης των σημείων που είναι χαρακτηρισμένα σαν "δεδομένα εκπαίδευσης".

Ένα νευρωνικό δίκτυο το οποίο είναι σχεδιασμένο ώστε να γενικεύει καλά, θα παράγει σωστή αντιστοίχιση εισόδου-εξόδου ακόμη και όταν η είσοδος του είναι ελαφρώς διαφορετική από τα παραδείγματα που χρησιμοποιήθηκαν για την εκπαίδευσή του. Ωστόσο όταν ένα δίκτυο μαθαίνει υπερβολικά πολλά παραδείγματα εισόδου-εξόδου, μπορεί να οδηγηθεί σε μια κατάσταση όπου απομνημονεύει τα δεδομένα εκπαίδευσης και λειτουργεί μόνο βάσει αυτών. Αυτό το φαινόμενο ονομάζεται υπερπροσαρμογή (overfitting) ή υπερεκπαίδευση (overtraining) του δικτύου. Ένα τέτοιο δίκτυο χάνει την δυνατότητα να γενικεύει βασιζόμενο σε παρόμοια πρότυπα εισόδου-εξόδου [5,6,8].



**Σχήμα 2.27** (α) Έξοδος δικτύου το οποίο γενικεύει σωστά και (β) έξοδος δικτύου που έχει υπερεκαπιδευτεί [5].

**ΤΟ ΘΕΩΡΗΜΑ ΚΑΘΟΛΙΚΗΣ ΠΡΟΣΕΓΓΙΣΗΣ**

Για ένα perceptron πολλαπλών επιπέδων από τη σκοπιά της αντιστοίχισης εισόδου-εξόδου, το θεώρημα καθολικής προσέγγισης εξασφαλίζει ότι με ένα κρυμμένο επίπεδο η προσέγγιση μπορεί να έχει οποιαδήποτε ακρίβεια επιθυμούμε. Αυτό φυσικά δε σημαίνει ότι το δίκτυο αυτό είναι το βέλτιστο από άποψη χρόνου εκπαίδευσης ή υλοποίησης [5].

## ΚΕΦΑΛΑΙΟ 3<sup>ο</sup> : ΧΡΗΣΗ ΤΝΔ ΓΙΑ ΤΗ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ ΚΑΤΕΡΓΑΣΙΩΝ

### 3.1 Εισαγωγή

Τα τεχνητά νευρωνικά δίκτυα βρίσκουν χρήση σε διάφορα πεδία ένα εκ των οποίων είναι η βελτιστοποίηση κατεργασιών όπως για παράδειγμα η εξαγωγή συμπερασμάτων για την επιρροή των διαφόρων παραμέτρων μιας κατεργασίας στο τελικό αποτέλεσμα. Στη παρούσα εργασία μελετάμε τα νευρωνικά δίκτυα ως προς τη δυνατότητά τους να προβλέψουν την επιφανειακή τραχύτητα επιφάνειας που προέκυψε από τόννευση σε υλικό υψηλής σκληρότητας. Όσον αφορά τη πρόβλεψη της επιφανειακής τραχύτητας, τα ΤΝΔ έχουν μελετηθεί σε μεγάλο βαθμό με καταγεγραμμένες εργασίες να αναφέρονται σε κατεργασίες φρεζαρίσματος, τόννευσης κλπ. Παρακάτω θα αναφερθούμε περιληπτικά σε κάποιες από αυτές.

### 3.2 Εργασίες στη χρήση των ΤΝΔ για τη βελτιστοποίηση κατεργασιών

Στην εργασία τους οι Anuja Beatricea et al. [9] αναφέρουν ότι από τη διεθνή βιβλιογραφία, παρατήρησαν πως διάφορα ΤΝΔ είχαν χρησιμοποιηθεί επιτυχώς από ερευνητές για την πρόβλεψη της επιφανειακής τραχύτητας σε κατεργασία με τόννο σε σκληρά υλικά, αλλά δεν υπήρχε κάποια εργασία στην πρόβλεψη επιφανειακής τραχύτητας με την χρήση παραμέτρων κοπής κατά την τόννευση του σκληρού υλικού AISI H13 με την ελάχιστη εφαρμογή υγρού κοπής. Έτσι πραγματοποιήθηκε από μέρους τους μια προσπάθεια να αναπτυχθεί ένα μοντέλο βασισμένο στα ΤΝΔ το οποίο να πραγματοποιεί αυτήν ακριβώς τη διαδικασία. Κατά την έρευνά τους εκπαιδεύτηκαν δίκτυα με διαφορετικές αρχιτεκτονικές για ένα συγκεκριμένο αριθμό εποχών και έπειτα ελέγησαν-δοκιμάστηκαν για ένα συγκεκριμένο αριθμό διαφορετικών παραδειγμάτων εισόδου-εξόδου από αυτά της εκπαίδευσης.

Αυτή η έρευνα εφαρμόστηκε σε συνολικά 60 αρχιτεκτονικές από τις οποίες οι 8 αρχιτεκτονικές 3-1-1, 3-3-1, 3-6-1, 3-7-1, 3-1-1-1, 3-3-3-1, 3-6-6-1, 3-7-7-1 ακολουθούν τον κανόνα του Zhang ο οποίος προτείνει ο αριθμός νευρώνων στα κρυφά επίπεδα να είναι  $n/2$ ,  $n$ ,  $2n$ ,  $2n+1$  όπου  $n$  ο αριθμός των νευρώνων στο επίπεδο εισόδου.

Στη συγκεκριμένη εργασία τα δίκτυα κατασκευάστηκαν σε Matlab και η συνάρτηση μάθησης ήταν η 'learngdm' ενώ η συνάρτηση εκπαίδευσης ήταν η 'trainlm'. Η συνάρτηση ενεργοποίησης για τους νευρώνες

των ΤΝΔ ήταν η "tansig":  $f(x) = \frac{1}{1 + e^{-x}}$

Τα κρυφά επίπεδα των δικτύων μεταβάλλονταν βαθμιαία σε μονοστρωματικούς και πολυστρωματικούς τύπους δικτύων με διαφορετικούς αριθμούς νευρώνων να χρησιμοποιούν τη μέθοδο διόρθωσης σφάλματος για την καλύτερη αρχιτεκτονική με το μικρότερο MSE σφάλμα.

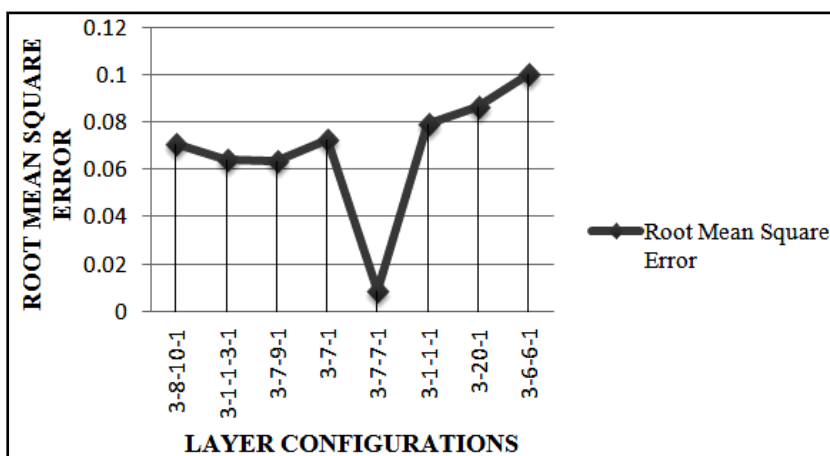
Από τα 27 παραδείγματα του πίνακα 3.1 ο οποίος παρατίθεται στην εργασία τους, χρησιμοποιήθηκαν τα 23 για την εκπαίδευση του δικτύου και τα 4 υπόλοιπα για τον έλεγχο, οπότε ο λόγος ήταν 85% :15% μεταξύ του αριθμού παραδειγμάτων εκπαίδευσης και των παραδειγμάτων ελέγχου.

S. No	f(mm/rev)	v(m/min)	d(mm)	Ra(μm)	Train/Test
1	0.05	75	0.5	1.28	Training
2	0.05	75	0.75	1.1	Training
3	0.05	75	1	1.58	Training
4	0.05	95	0.5	0.91	Training
5	0.05	95	0.75	1.03	Training
6	0.05	95	1	1.13	Training
7	0.05	115	0.5	1.38	Training
8	0.05	115	0.75	1.34	Training
9	0.05	115	1	1.45	Test
10	0.075	75	0.5	1.19	Training

11	0.075	75	0.75	1.09	Training
12	0.075	75	1	1.48	Training
13	0.075	95	0.5	1.45	Training
14	0.075	95	0.75	1.86	Training
15	0.075	95	1	1.82	Test
16	0.075	115	0.5	1.43	Training
17	0.075	115	0.75	1.7	Training
18	0.075	115	1	1.65	Training
19	0.1	75	0.5	1.64	Test
20	0.1	75	0.75	1.13	Training
21	0.1	75	1	1.78	Training
22	0.1	95	0.5	1.77	Training
23	0.1	95	0.75	1.83	Training
24	0.1	95	1	1.53	Training
25	0.1	115	0.5	1.32	Test
26	0.1	115	0.75	1.6	Training
27	0.1	115	1	1.79	Training

**Πίνακας 3.1** Πειραματικά δεδομένα που συλλέχθηκαν από 27 διαφορετικά πειράματα [9].

Παρακάτω παρατίθενται τα αποτελέσματα των σφαλμάτων MSE για κάποιες από τις αρχιτεκτονικές που μελετήθηκαν στην συγκεκριμένη εργασία:



**Σχήμα.3.1** Τα σφάλματα MSE για τις διάφορες αρχιτεκτονικές ΤΝΔ [9].

Η αρχιτεκτονική με 3 νευρώνες στο επίπεδο εισόδου, 2 κρυφά επίπεδα με 7 νευρώνες το καθένα και ένα νευρώνα στο επίπεδο εξόδου (3-7-7-1) αποδείχθηκε να έχει τη μικρότερη τιμή MSE: 0.008 έπειτα από 10000 εποχές εκπαίδευσης.

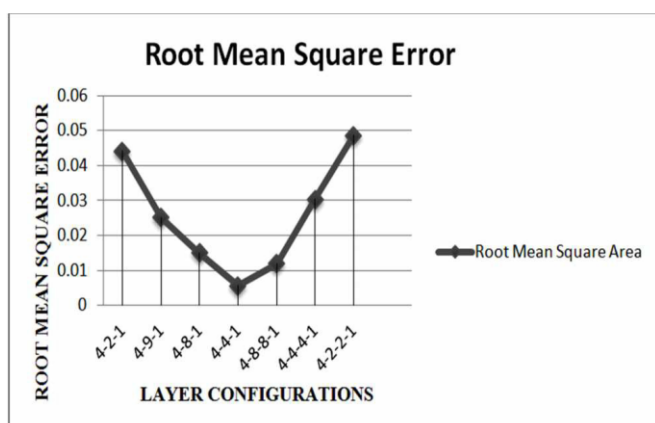
Όπως φαίνεται στον παρακάτω πίνακα το σφάλμα στην πρόβλεψη της επιφανειακής τραχύτητας από τα ΤΝΔ κυμαίνεται κάτω από 10%.

Testing Data				Επιφανειακή Τραχύτητα		
No	f	v	d	Πειραματικά αποτελέσματα	Πρόβλεψη από το ΤΝΔ	% Σφάλμα
1	0.05	115	1	1.45	1.35	6.89
2	0.075	95	1	1.82	1.74	4.39
3	0.1	75	0.5	1.64	1.58	3.65
4	0.1	115	0.5	1.32	1.26	4.54

**Πίνακας 3.2** Αποτελέσματα για τα 4 σετ ελέγχου [9].

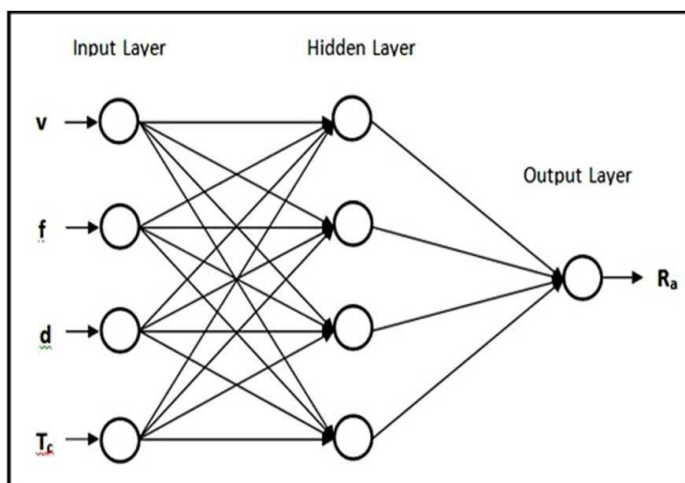
Στην εργασία των Agulraj et al. [10], έγινε προσπάθεια να εισαχθεί η παράμετρος της θερμοκρασίας μαζί με τις παραμέτρους κοπής για την πρόβλεψη της επιφανειακής τραχύτητας κατά την τόννευση του χάλυβα H13. Το ΤΝΔ που κατασκευάστηκε με εισόδους τις παραμέτρους κοπής της προαναφερθείσας εργασίας συγκρίθηκε με ένα νέο πανομοιότυπο ΤΝΔ που τώρα όμως εμπεριείχε και τον παράγοντα της θερμοκρασίας κοπής. Έπειτα αναλύθηκαν και τα δύο ως προς την ικανότητά τους να προβλέψουν την επιφανειακή τραχύτητα ( $R_a$ ). Τα παραδείγματα για την εκπαίδευση και τον έλεγχο του δικτύου ήταν τα ίδια με του προηγούμενου δικτύου και φαίνονται στον πίνακα 3.1. Όπως φαίνεται και παρακάτω, το μοντέλο με την θερμοκρασιακή παράμετρο αποδείχθηκε ανώτερο του προηγούμενου, καθώς υπάρχουν αρκετές διαφορετικές αρχιτεκτονικές με μικρότερο MSE.

Από τις διαφορετικές αρχιτεκτονικές αυτή με το μικρότερο MSE τιμή ήταν το 4-4-1, με MSE σφάλμα 0.005 στους 10000 κύκλους (εποχές).



Σχήμα 3.2 Τα σφάλματα MSE για τις διαφορετικές αρχιτεκτονικές ΤΝΔ [10].

Η αρχιτεκτονική 4-4-1 με το μικρότερο σφάλμα MSE φαίνεται και στο παρακάτω σχήμα:



Σχήμα 3.4 Η Αρχιτεκτονική του βέλτιστου δικτύου 4-4-1 (στο επίπεδο εισόδου συνυπολογίζεται ο όρος  $T_c$ ) [10].

Όπως φαίνεται στον παρακάτω πίνακα το σφάλμα στην πρόβλεψη της επιφανειακής τραχύτητας από τα ΤΝΔ που συνυπολογίζουν στις εισόδους τους την παράμετρο της θερμοκρασίας κατά την κοπή είναι σαφώς βελτιωμένο συγκριτικά με την προηγούμενη εργασία.

Testing Data				Επιφανειακή τραχύτητα		
S.No	f	v	d	Πειραματικά αποτελέσματα	Πρόβλεψη από τα ΤΝΔ με την παράμετρο θερμοκρασίας	% Σφάλμα
1	0.05	115	1	1.45	1.411	2.6
2	0.075	95	1	1.82	1.761	3.2
3	0.1	75	0.5	1.64	1.615	1.5
4	0.1	115	0.5	1.32	1.28	2.98

**Πίνακας 3.3** Αποτελέσματα για τα 4 σετ ελέγχου. [10].

Στην εργασία των Aouici et al. [11], έγινε μια πειραματική δουλειά η οποία εστιάζει στη πρόβλεψη και τη βελτιστοποίηση και της επιφανειακής τραχύτητας και των συνιστωσών που διενεργούν στην δύναμη κοπής κατά την τόννευση του χάλυβα AISI H11 με κοπτικό εργαλείο CBN. Τα πειράματα τόννευσης πραγματοποιήθηκαν χωρίς τη χρήση υγρού κοπής. Το υλικό του κομματιού ήταν το AISI H1, το οποίο χρησιμοποιείται ευρέως σε εν θερμώ διαμορφώσεις. Το CBN κοπτικό εργαλείο που χρησιμοποιήθηκε στα πειράματα ήταν το CBN 7020. Οι τρεις συνιστώσες των δυνάμεων κοπής (feed force ( $F_a$ ), thrust force ( $F_r$ ) και tangential force ( $F_v$ )), καταγράφηκαν πειραματικά με δυναμόμετρο.

Η μέθοδος RSM (response surface methodology) είναι ικανή να βρει τη σχέση μεταξύ ανεξάρτητων μεταβλητών-εισόδων μιας διαδικασίας και των δεδομένων εξόδου. Στη συγκεκριμένη περίπτωση η σχέση μεταξύ των εισόδων που είναι οι συνθήκες κοπής (ταχύτητα κοπής ( $V_c$ ), πρόωση ( $f$ ), βάθος κοπής ( $a_p$ ) και σκληρότητα του κομματιού ( $H$ )) και της εξόδου  $Y$  η οποία είναι η επιφανειακή τραχύτητα ( $Ra$ ) ή οι συνιστώσες της δύναμης κοπής ( $F_a$ ,  $F_r$  and  $F_v$ )) δίνεται από:

$$Y = \varphi(V_c, f, a_p, H)$$

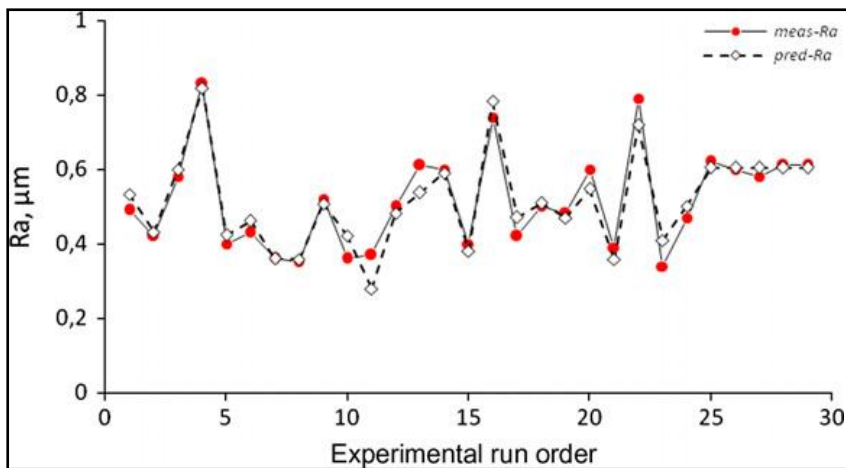
όπου  $\varphi$  είναι η συνάρτηση της απόκρισης.

Το μοντέλο για την τραχύτητα  $Ra$  δίνεται παρακάτω από την εξίσωση:

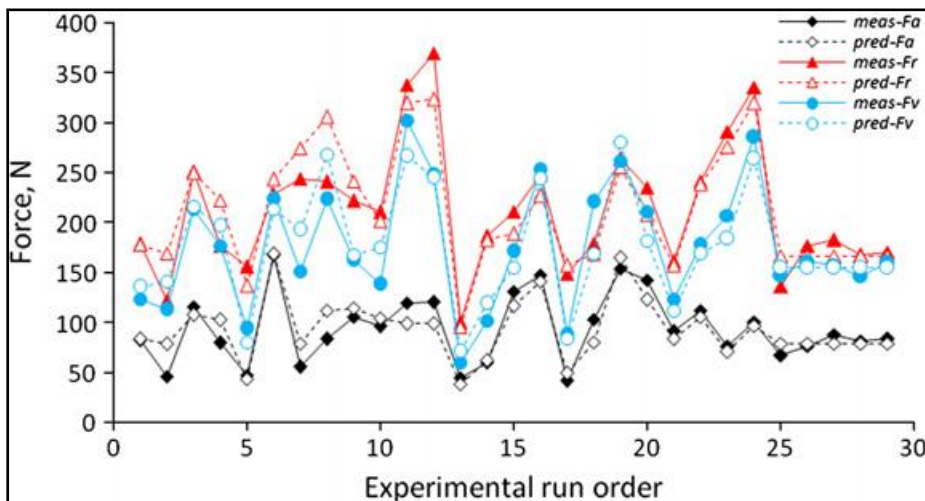
$$\begin{aligned} Ra = & -9.872 - 0.016 \cdot V_c + 2.75 \cdot f + 0.4908 \cdot a_p \\ & + 0.496 \cdot H + 0.0334 V_c \cdot f + 1.111 \cdot 10^{-3} V_c \cdot a_p \\ & + 2.416 \cdot 10^{-4} \cdot V_c \cdot H + 14.583 \cdot f \cdot a_p - 0.0337 \cdot f \cdot H \\ & - 0.01334 a_p \cdot H - 1.14310^{-5} \cdot V_c^2 + 20.364 \cdot f^2 \\ & - 2.829 \cdot a_p^2 - 5.596 \cdot H^2 \end{aligned}$$

Αντίστοιχα αναπτύχθηκαν και τα μοντέλα παλινδρόμησης για τις συνιστώσες της δύναμης κοπής  $F_a$ ,  $F_r$ ,  $F_v$ . Οι τιμές που προβλέφθηκαν για τις αποκρίσεις που απεικονίζουν την επιφανειακή τραχύτητα και τις συνιστώσες της δύναμης κοπής ( $F_a$ ,  $F_r$  και  $F_v$ ), από τις εξισώσεις παλινδρόμησης που αντιστοιχούν στους διαφορετικούς συνδυασμούς των παραμέτρων της κατεργασίας συγκρίνονται με τις πειραματικές μετρήσεις και απεικονίζονται στα παρακάτω σχήματα:





Σχήμα 3.5 Σύγκριση μεταξύ πειραματικών και προβλεπόμενων τιμών για την επιφανειακή τραχύτητα Ra [11].



Σχήμα 3.6 Σύγκριση μεταξύ πειραματικών και προβλεπόμενων τιμών για τις συνιστώσες της δύναμης κοπής [11].

Το συμπέρασμα της εργασίας τους είναι ότι, η εφαρμογή της μεθόδου RSM στην τόνρευση του χάλυβα AISI H11 με κοπτικό εργαλείο από CBN7020 οδήγησε επιτυχώς στην απόκτηση μαθηματικών μοντέλων τόσο για την επιφανειακή τραχύτητα όσο και για τις συνιστώσες της δύναμης κοπής. Στη συνέχεια εξετάστηκαν και υπολογίστηκαν οι βέλτιστες τιμές για τις παραμέτρους της κατεργασίας και καταλήξαν στα εξής:

1. Η ανάλυση των παραμέτρων κατεργασίας με τη χρήση της μεθόδου RSM επιτρέπει στη επιθεώρηση της κάθε μιας ξεχωριστά ως προς την επιρροή της στην απόκριση της κατεργασίας (όπως η επιφανειακή τραχύτητα και οι συνιστώσες της δύναμης κοπής)
2. Η σύγκριση των πειραματικών και των τιμών που προβλέφθηκαν για την επιφανειακή τραχύτητα και τις συνιστώσες της δύναμης κοπής δείχνουν ότι επιτεύχθηκε η εύρεση μιας καλής σχέσης μεταξύ τους.
3. Επιπλέον, η ταχύτητα πρόωσης και η σκληρότητα του κατεργαζόμενου κομματιού έχουν σημαντική στατιστικά επιρροή στην επιφανειακή τραχύτητα.
4. Η δύναμη πρόωσης ( $F_w$ ) και η δύναμη κοπής ( $F_v$ ) επηρεάζονται σημαντικά από το βάθος κοπής, (56.77%) and (31.50%) αντίστοιχα. Αντίθετα, η ταχύτητα κοπής έχει μικρή επιρροή (0.14%).
5. Η βέλτιστη επιφανειακή τραχύτητα επιτεύχθηκε για χαμηλότερες προώσεις και υψηλότερες ταχύτητες κοπής

Στην εργασία των Ozel και Karpat [12] επιθεωρήθηκαν πειραματικά η επιρροή της γεωμετρίας της αιχμής του κοπτικού εργαλείου, η σκληρότητα του κατεργαζόμενου κομματιού, η πρόωση και η ταχύτητα κοπής για την επιφανειακή τραχύτητα και τη φθορά του κοπτικού εργαλείου κατά την τόννευση χωρίς υγρό κοπής του χάλυβα AISI H13.

Για τα πειράματα χρησιμοποιήθηκαν κοπτικά με χαμηλή περιεκτικότητα σε CBN με δυο διαφορετικές μορφές κοπτικής ακμής και μπάρες από το σκληρό χάλυβα AISI H13.

Πραγματοποιήθηκε παραγοντικός σχεδιασμός πειραμάτων 4 παραγόντων και 2 επιπέδων για την διερεύνηση της επιρροής της γεωμετρίας της αιχμής του κοπτικού, της σκληρότητας του κομματιού, της πρόωσης και της ταχύτητας κοπής στην επιφανειακή τραχύτητα και τη φθορά του εργαλείου.

Έγιναν 16 εκτελέσεις του κάθε των συνδυασμών των επιπέδων με αποτέλεσμα να έχουμε συνολικά 256 τεστ. Η κατά μήκος τόννευση έγινε σε μπάρες του συγκεκριμένου υλικού σε έναν CNC τόρνο υψηλής ακρίβειας (Romí Centur 35E) με σταθερό βάθος κοπής στα 0.254 mm.

Το βασικό θεωρητικό μοντέλο της επιφανειακής τραχύτητα δίνεται από την εξίσωση:

$$Ra = \frac{f^2}{32r_e}$$

όπου  $f$  είναι η πρόωση και  $r_e$  είναι η ακτίνα της ακμής του εργαλείου κοπής.

Σύμφωνα με αυτό το μοντέλο, χρειάζεται να μειωθεί η πρόωση ή να αυξηθεί η ακτίνα της ακμής του κοπτικού για να βελτιωθεί η επιθυμητή επιφανειακή τραχύτητα. Ωστόσο, υπάρχουν διάφορα προβλήματα με αυτό το μοντέλο. Καταρχάς, δεν λαμβάνει υπόψιν οποιαδήποτε ατέλεια της κατεργασίας όπως τις δονήσεις του εργαλείου ή την επικόλληση αποβλίτου. Δεύτερον, υπάρχουν πρακτικοί περιορισμοί σε αυτό το μοντέλο καθώς συγκεκριμένα εργαλεία (όπως τα CBN) απαιτούν συγκεκριμένες γεωμετρίες για να βελτιωθεί η διάρκεια ζωής του κοπτικού εργαλείου.

Στη συγκεκριμένη εργασία προτείνεται ένα τροποποιημένο μοντέλο τόσο για την επιφανειακή τραχύτητα όσο και για τη φθορά του εργαλείου το οποίο λαμβάνει υπόψιν την τόννευση σκληρών υλικών με εργαλεία CBN.

Σε αυτό το μοντέλο, η επιφανειακή τραχύτητα ή το βάθος της φθοράς του κοπτικού εργαλείου είναι συνάρτηση της σκληρότητας του κομματιού, της περιεκτικότητας του κοπτικού σε CBN, της ακτίνας της ακμής του κοπτικού εργαλείου, της ταχύτητας κοπής, της πρόωσης και του χρόνου κοπής.

Σε αυτή την εργασία χρησιμοποιήθηκαν δύο διαφορετικά νευρωνικά δίκτυα.

**Στο πρώτο σύνολο** η επιφανειακή τραχύτητα και η φθορά του εργαλείου προβλέπονται ταυτόχρονα με ένα πρόσθιας τροφοδότησης και πολλών επιπέδων νευρωνικό δίκτυο όπως φαίνεται στο σχήμα 3.7a χρησιμοποιώντας ως εισόδους τις παρακάτω 5 παραμέτρους:

- γεωμετρία της ακμής του κοπτικού
- σκληρότητα του κατεργαζόμενου κομματιού σε Rockwell-C
- ταχύτητα κοπής (m/min)
- πρόωση (mm/rev)
- μήκος κοπής αξονικά (mm)

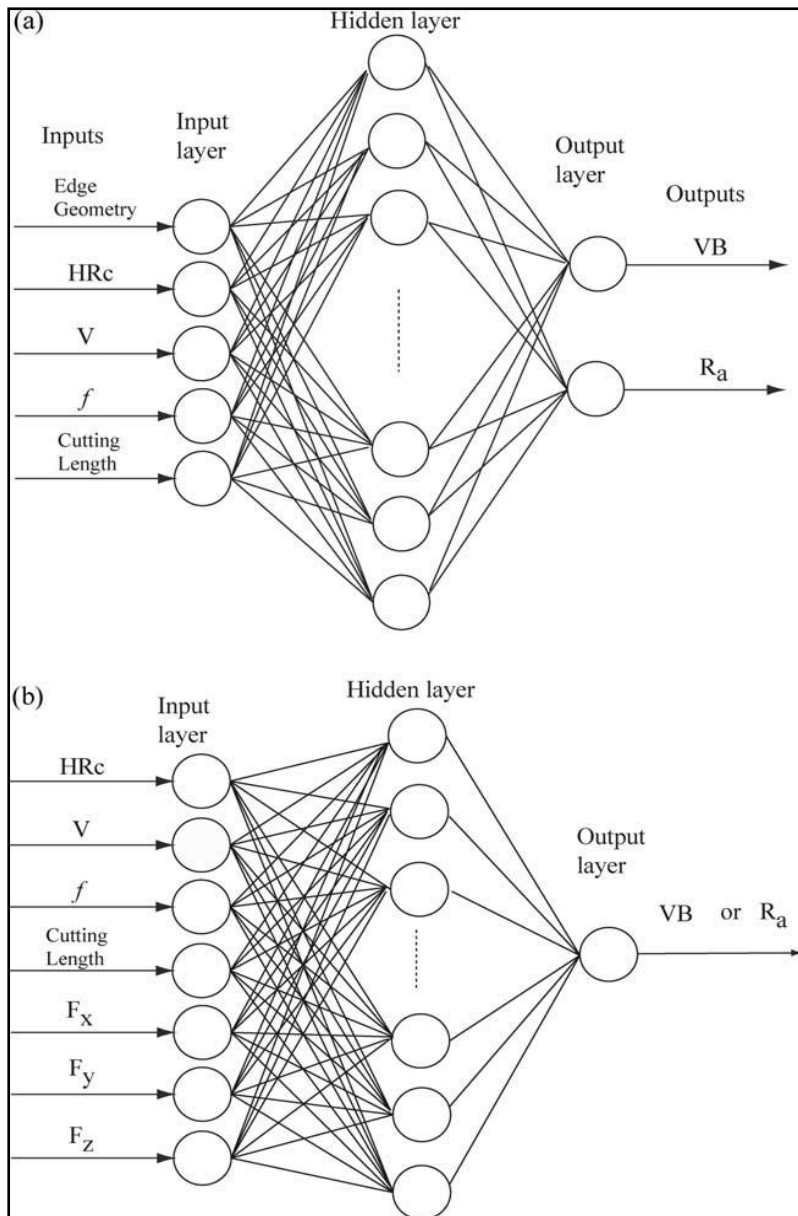
Αυτό το νευρωνικό δίκτυο εκπαιδεύτηκε με 173 παραδείγματα και ελέγχθηκε με 36 παραδείγματα τα οποία είχαν επιλεγεί τυχαία από διαφορετικά παραδείγματα. Συνολικά τα παραδείγματα ήταν 209.

**Στο δεύτερο σύνολο**, σχεδιάστηκαν νευρωνικά δίκτυα για κοπτικές αιχμές chamfered και honed ξεχωριστά. Έχει αναφερθεί ότι τα σήματα των δυνάμεων κοπής είναι ευαίσθητα στη φθορά του εργαλείου. Με βάση την αξιοπιστία των μετρήσεων των δυνάμεων κοπής, ως είσοδοι επιλέχθηκαν οι μέσες τιμές τους (βλ. σχήμα 3.7b) ώστε η πρόβλεψη της επιφανειακής τραχύτητας και της φθοράς του κοπτικού να είναι καλύτερη. Τα δίκτυα αυτά θα είχαν μια έξοδο το καθένα. Με αυτό το τρόπο, μειώθηκε το μέγεθος των δικτύων οπότε επιτεύχθηκε γρηγορότερη σύγκλιση και καλύτερη πρόβλεψη. Τελικά δημιουργήθηκαν 4 διαφορετικά ΤΝΔ:

1. chamfered αιχμή - επιφανειακή τραχύτητα
2. chamfered αιχμή - φθορά εργαλείου
3. honed αιχμή - επιφανειακή τραχύτητα
4. honed αιχμή - φθορά εργαλείου

με 7 εισόδους και μια έξοδο το καθένα όπου οι είσοδοι είναι:

- σκληρότητα του κατεργαζόμενου κομματιού σε Rockwell-C
- ταχύτητα κοπής (m/min)
- πρόωση (mm/rev),
- μήκος κοπής αξονικά (mm)
- 3 μέσες τιμές για τις δυνάμεις  $F_x$ ,  $F_y$ ,  $F_z$  (N) που μετρήθηκαν πειραματικά.



**Σχήμα 3.7** Τα ΤΝΔ για τη πρόβλεψη της επιφανειακής τραχύτητας και της φθοράς του κοπτικού εργαλείου [12].

Αυτά τα ΤΝΔ εκπαιδεύτηκαν με 111 παραδείγματα και ελέγχθηκαν με 16 παραδείγματα.

Επειτά επιλέχθηκε η δομή ΤΝΔ 7–8–1 για την πρόβλεψη της φθοράς του κοπτικού με chamfered αιχμή και οι δομές 7–10–1, 7–10–1, και 7–13–1 για πρόβλεψη της φθοράς του κοπτικού με honed αιχμή.

Τέλος, έγινε σύγκριση μεταξύ των προβλέψεων της επιφανειακής τραχύτητας και της φθοράς του κοπτικού και των μοντέλων που είναι βασισμένα σε παλινδρόμηση και των ΤΝΔ με νέα πειραματικά δεδομένα που δεν δόθηκαν στην εκπαίδευση των ΤΝΔ.

Τα συμπεράσματα από την συγκεκριμένη εργασία είναι ότι τα συγκεκριμένα ΤΝΔ ήταν ικανά να προβλέψουν την επιφανειακή τραχύτητα και τη φθορά του εργαλείου εντός του εύρους παραδειγμάτων με τα οποία

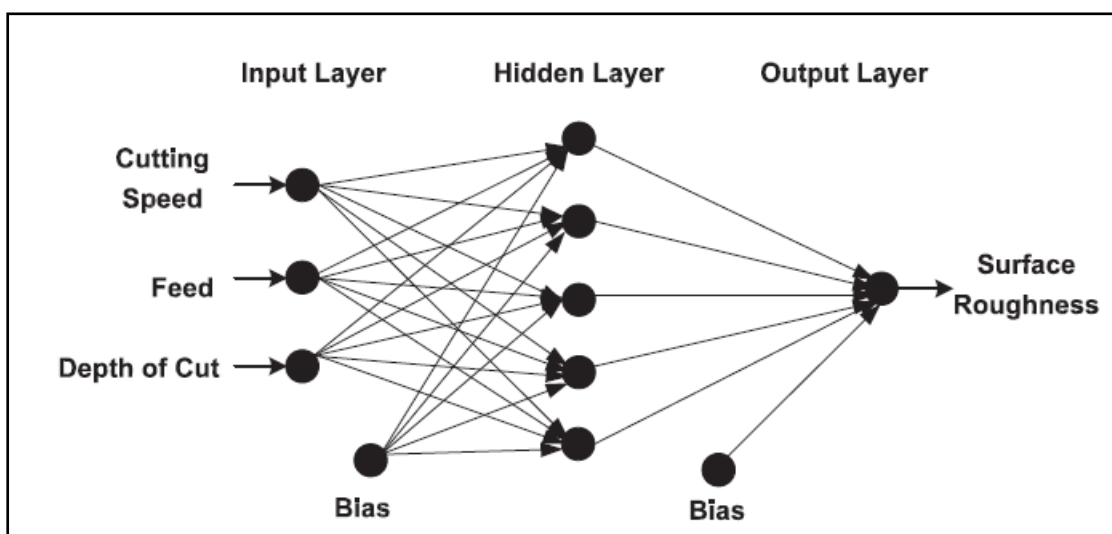
εκπαιδύτηκαν. Επίσης, συγκρινόμενα με τα μοντέλα παλινδρόμησης αποδείχθηκε ότι τα TND έδιναν καλύτερα αποτελέσματα επειδή γενικά μπορούν να προσεγγίσουν καλύτερα μη-γραμμικά μοντέλα σε σχέση με τα πρώτα.

Οι αλγόριθμοι εκπαίδευσης που χρησιμοποιήθηκαν ήταν οι Bayesian regularization με τον Levenberg – Marquardt. Τα TND με εισόδους τις δυνάμεις κοπής και μια μόνο έξοδο έδωσαν καλύτερα αποτελέσματα από αυτά με τις δυο εξόδους τα οποία προβλέπαν ταυτόχρονα την επιφανειακή τραχύτητα και τη φθορά του εργαλείου.

Η εργασία των Asiltórk και Hunkas [13] είχε ως στόχο να αναπτυχθεί μια μέθοδος βασισμένη στα TND και τα μοντέλα πολλαπλής παλινδρόμησης για να προβλεφθεί η επιφανειακή τραχύτητα του χάλυβα AISI 1040 έπειτα από τόννευση. Έτσι, πραγματοποιήθηκαν πειράματα ώστε να διερευνηθεί η επιρροή των παραμέτρων κοπής (όπως η ταχύτητα κοπής, η πρόωση, και το βάθος κοπής) στην επιφανειακή τραχύτητα. Τα μοντέλα πολλαπλής παλινδρόμησης ελέγχθηκαν με τη μέθοδο ANOVA. Αναπτύχθηκαν αρχιτεκτονικές MLP με τον αλγόριθμο οπισθοδιάδοσης σφάλματος και τελικά η απόδοση των μοντέλων πολλαπλής παλινδρόμησης και των μοντέλων βασισμένα στα TND συγκρίθηκαν με στατιστικές μεθόδους.

Η πολλαπλή παλινδρόμηση είναι μια τεχνική που χρησιμοποιείται στη στατιστική και μας επιτρέπει να βρούμε μια σχέση μεταξύ μίας ή περισσότερων συνεχών ή διακριτών ανεξάρτητων μεταβλητών και μιας συνεχούς εξαρτημένης μεταβλητής. Μπορεί να χρησιμοποιηθεί για μια ποικιλία προβλημάτων όπως η ανάλυση πειραματικών δεδομένων. Επομένως, θα μπορούσε να βοηθήσει στην πρόβλεψη της επιφανειακής τραχύτητας.

Σε αυτή την εργασία, το TND που χρησιμοποιήθηκε για την πρόβλεψη της επιφανειακής τραχύτητας έπειτα από τόννευση φαίνεται στο σχήμα 3.8. Το συγκεκριμένο TND έχει ένα επίπεδο εισόδου, ένα κρυμμένο επίπεδο και ένα επίπεδο εξόδου. Χρησιμοποιήθηκαν οι αλγόριθμοι οπισθοδιάδοσης σφάλματος όπως ο scaled conjugate gradient (SCG) και ο Levenberg–Marquardt (LM) για την ανανέωση των παραμέτρων στα κρυμμένα επίπεδα.



Σχήμα 3.8 Η δομή του TND που χρησιμοποιήθηκε [13].

Η ταχύτητα κοπής, η πρόωση και το βάθος κοπής θεωρήθηκαν οι παράμετροι της κατεργασίας που θα τροφοδοτούσαν το επίπεδο εισόδου του TND το οποίο επομένως αποτελείται από 3 νευρώνες ενώ το επίπεδο εξόδου από έναν ο οποίος αποδίδει την προβλεπόμενη τιμή της επιφανειακής τραχύτητας. Παρατηρήθηκε ότι οι συναρτήσεις ενεργοποίησης, logsig και tansig, παρέχουν την ίδια περίπου απόδοση στο TND σε διαφορετικά προβλήματα οπότε τελικά χρησιμοποιήθηκε μόνο η logsig. Η μέθοδος δοκιμής - έλεγχος σφάλματος χρησιμοποιήθηκε για να αποφασιστεί ο αριθμός νευρώνων σε κάθε επίπεδο του TND. Ο αριθμός των κρυφών νευρώνων αποφασίστηκε να είναι 4 και 5 νευρώνες. Καθώς οι είσοδοι βρίσκονταν σε διαφορετικό εύρος τιμών, κανονικοποιήθηκαν σε ένα εύρος μεταξύ 0.1–0.9 ώστε οι νευρώνες να μη φθάνουν

τη συνάρτηση ενεργοποίησης στα όριά της. Ο μέγιστος αριθμός εποχών και ο ρυθμός μάθησης επιλέχθηκαν να είναι 10.000 και 0.9 αντίστοιχα.

Τα παραδείγματα προέρχονται από πειράματα που πραγματοποιήθηκαν σε τόρνο CNC ενώ τα δοκίμια που χρησιμοποιήθηκαν ήταν από το χάλυβα AISI 1040 το οποίο έχει σκληρυνθεί στα 35 HRC. Για να παραμείνουν οι συνθήκες κοπής ίδιες, το κάθε πείραμα έγινε με νέο κοπτικό εργαλείο. Τα πειράματα που έγιναν ήταν συνολικά 27 και πραγματοποιήθηκαν χωρίς ψυκτικό. Οι παράμετροι της τραχύτητας γενικά εξαρτώνται από της συνθήκες της κατεργασίας όπως η πρόωση, το βάθος κοπής, η ταχύτητα κοπής το κοπτικό εργαλείο κ.α. Στη συγκεκριμένη εργασία επιλέχθηκαν 3 από αυτές τις παραμέτρους να είναι η βάση για την ανάλυση και αυτές ήταν η ταχύτητα κοπής, η πρόωση και το βάθος κοπής.

Τα αποτελέσματα έδειξαν πως η πρόωση είναι ο κυρίαρχος παράγοντας που επηρεάζει την επιφανειακή τραχύτητα, και ακολουθούν το βάθος κοπής και η ταχύτητα κοπής. Από τους αλγόριθμους οπισθοδιάδοσης σφάλματος (SCG και LM), που χρησιμοποιήθηκαν για την εκπαίδευση των δικτύων, τα καλύτερα αποτελέσματα με το μικρότερο σφάλμα τα έδωσε ο αλγόριθμος SCG με 5 κρυφούς νευρώνες. Οι τιμές που προβλέφθηκαν για την επιφανειακή τραχύτητα ήταν πολύ κοντά στις πειραματικές οπότε μπορεί να ειπωθεί ότι τα συγκεκριμένα μοντέλα μπορούν να προβλέψουν αποτελεσματικά την επιφανειακή τραχύτητα του κομματιού έπειτα από τόννευση.

Οι παραπάνω εργασίες που παρουσιάσαμε μας δίνουν μια γενική εικόνα για τη χρήση των τεχνητών νευρωνικών δικτύων στη πρόβλεψη της επιφανειακής τραχύτητας, που είναι και ζητούμενο στην παρούσα εργασία, καθώς επίσης και μια ιδέα για άλλες μεθόδους πρόβλεψής της. Μπορούμε να συμπεράνουμε ότι η μέθοδος της κανονικοποίησης των δεδομένων πριν την εκπαίδευση των ΤΝΔ καθώς και η χρήση περισσότερων του ενός κρυφών επιπέδων θα βοηθήσουν στην δημιουργία περισσότερο αξιόπιστων ΤΝΔ. Ακόμη, επιβεβαιώνεται ότι όσον αφορά τις μεθόδους εκπαίδευσης, γενικά χρησιμοποιούνται συχνά οι μέθοδοι scaled conjugate gradient, Levenberg–Marquardt. Τέλος, παρατηρούμε ότι γενικά γίνεται διαχωρισμός των δεδομένων σε training και test sets αλλά δεν χρησιμοποιείται validation set για τον έλεγχο της εκπαίδευσης, το οποίο set όμως θα χρησιμοποιηθεί στην παρούσα εργασία όπως θα δούμε παρακάτω.

## **Β' ΠΕΙΡΑΜΑΤΙΚΟ ΜΕΡΟΣ**

## ΚΕΦΑΛΑΙΟ 4<sup>ο</sup> ΠΕΙΡΑΜΑΤΙΚΟ ΜΕΡΟΣ

### 4.1 Εισαγωγή

Στην παρούσα εργασία προσπαθήσαμε να συγκρίνουμε την απόδοση των τεχνητών νευρωνικών δικτύων με την απόδοση των μοντέλων που προτείνονται στην εργασία των Agrawal et al. [1]. Στα μοντέλα αυτά χρησιμοποιήθηκαν πειραματικά δεδομένα τα οποία αποτέλεσαν και δεδομένα εκπαίδευσης των νευρωνικών δικτύων που αναπτύξαμε εμείς. Στην δημοσιευμένη αυτή εργασία αναφέρεται ότι ένα από τα τρία μοντέλα που αναπτύχθηκαν-εξετάστηκαν, έχει αποδειχθεί να είναι ανώτερο πολλών άλλων υπολογιστικών μεθόδων μια από τις οποίες είναι και τα νευρωνικά δίκτυα. Κατά το τέλος της εργασίας θα επικεντρωθούμε στο συγκεκριμένο μοντέλο που πράγματι υπερτερεί των δυο άλλων ώστε να το συγκρίνουμε με το βέλτιστο ΤΝΔ που αναπτύξαμε.

### 4.2 Περιγραφή της εργασίας προς σύγκριση

Στην εργασία τους οι Agrawal, Goel, Bin Rashid και Price προσπαθούν να διερευνήσουν την επιρροή των διαφόρων παραμέτρων κατεργασιών στην επιφανειακή τραχύτητα του κομματιού. Με πειραματικά αποτελέσματα από 39 торνεύσεις οι οποίες πραγματοποιήθηκαν σε CNC τόρνο Mori-Seiki SL-25Y (4-αξόνων) και με CBN κοπτικά εργαλεία (τύπου CNMA 12 04 08 S-B), μοντελοποίησαν τα δεδομένα αυτά με 3 διαφορετικούς μεθόδους παλινδρόμησης:

1. Το μοντέλο πολλαπλής παλινδρόμησης.
2. Την τεχνική quantile παλινδρόμησης (η οποία προσαρμόζει τις καμπύλες παλινδρόμησης σε άλλα σημεία της κατανομής της εξόδου και όχι απλά στις μέσες τιμές).
3. Την παλινδρόμηση με τυχαία δάση (Random Forest Regression - RFR) η οποία ψάχνει να επιτύχει την μέγιστη ακρίβεια στα προβλεπόμενα αποτελέσματα.

Στον πίνακα 4.1 παρουσιάζονται τα αποτελέσματα της επιφανειακής τραχύτητας για διάφορους συνδιασμούς πρόωσης ( $f$ ), βάθους κοπής ( $a_p$ ) και ταχύτητας κοπής ( $n$ ).

Πείραμα #i	Πρόωση ( $f$ ) (mm/rev)	Βάθος κοπής ( $a_p$ ) (mm)	Ταχύτητα κοπής ( $n$ ) (rpm)	Πειραματικές μετρήσεις τραχύτητας ( $R_a$ ) ( $\mu\text{m}$ )
1	0.08	0,1	1608	0,502
2	0.08	0,105	1250	0,532
3	0.08	0,2	858	0,5902
4	0.08	0,2	965	0,539
5	0.08	0,452	1850	0,592
6	0.08	0,542	1072	0,5693
7	0.08	0,935	1072	0,5821
8	0.09	0,083	2145	0,667
9	0.09	0,125	1000	0,735
10	0.09	0,144	1072	0,683
11	0.09	0,2	858	0,6776
12	0.09	0,2	965	0,6179
13	0.09	0,2	1072	0,742
14	0.09	0,542	965	0,718
15	0.09	0,542	1072	0,65
16	0.09	0,753	2050	0,764
17	0.09	0,935	1072	0,625
18	0.1	0,045	2145	0,77
19	0.1	0,048	2681	0,781
20	0.1	0,133	1608	0,773
21	0.1	0,2	858	0,6687

22	0.1	0,2	965	0,7029
23	0.1	0,234	2145	0,772
24	0.1	0,352	2220	0,784
25	0.1	0,542	1072	0,6769
26	0.1	0,558	1400	0,812
27	0.1	0,754	858	0,809
28	0.1	0,935	1072	0,6966
29	0.15	0,019	2681	1,251
30	0.15	0,06	1287	1,361
31	0.15	0,1	2681	1,193
32	0.15	0,2	858	1,134
33	0.15	0,2	965	1,0854
34	0.15	0,2	1072	1,316
35	0.15	0,278	1608	1,312
36	0.15	0,542	1072	1,1083
37	0.15	0,657	1600	1,345
38	0.15	0,906	2600	1,523
39	0.15	0,935	1072	1,1337

**Πίνακας 4.1** Τα αποτελέσματα της επιφανειακής τραχύτητας για διάφορους συνδιασμούς πρόωσης (f), βάθους κοπής ( $a_p$ ) και ταχύτητας κοπής (n) [1].

#### Το Μοντέλο Πολλαπλής Παλινδρόμησης

Στο μοντέλο πολλαπλής παλινδρόμησης, θεωρήθηκε η μέση τιμή της επιφανειακής τραχύτητας ως συνάρτηση τριών γραμμικών μέσων πρόβλεψης: της πρόωσης, του βάθους κοπής και της ταχύτητας κοπής θεωρώντας την παρακάτω γραμμική σχέση:

$$Ra_i = a + \beta_1 f_i + \beta_2 a_{pi} + \beta_3 n_i + \varepsilon_i$$

Η παραπάνω εξίσωση ορίζει μια ευθεία γραμμή. Η παράμετρος  $a$  είναι μια σταθερά και η παράμετρος  $\varepsilon_i$  αναπαριστά το σφάλμα υπολογισμού του μοντέλου. Οι παράμετροι  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$  αναπαριστούν την προσδοκώμενη αύξηση της απόκρισης  $Ra_i$  ανά ονάδα αλλαγής των  $f_i$ ,  $a_{pi}$  και  $n_i$  αντίστοιχα. Το μοντέλο A πο φαίνεται και στον πίνακα 4.2 δείχνει τα αποτελέσματα του μοντέλου πολλαπλής παλινδρόμησης που ορίζεται από την εξίσωση:

$$Ra_i = -0.279 + 9.455 f_i + 0.0539 a_{pi} + 5.61 \cdot 10^{-5} n_i$$

Όπως φαίνεται η πρόωση είναι ο πιο σημαντικός παράγοντας πρόβλεψης της επιφανειακής τραχύτητας.

Η παραπάνω εξίσωση προϋποθέτει ότι η σχέση μεταξύ της εξαρτημένης μεταβλητής  $Ra_i$  και των ανεξάρτητων μεταβλητών  $f_i$ ,  $a_{pi}$  και  $n_i$  είναι αθροιστική. Ωστόσο, η στιγμιαία επιρροή δυο ανεξάρτητων μεταβλητών μαζί μπορεί να μην είναι αθροιστική. Για παράδειγμα, η επίδραση της πρόωσης πιθανώς να εξαρτάται από το βάθος κοπής. Τέτοιου είδους φαινόμενα είναι γνωστά ως "φαινόμενα αλληλεπίδρασης" (interaction effects). Η εξίσωση πιο πάνω μπορεί να τροποποιηθεί έτσι ώστε να περιλαμβάνει τις αλληλεπιδράσεις κάθε ζεύγους ανεξάρτητων μεταβλητών και να πάρουμε τελικά την παρακάτω σχέση:

$$Ra_i = a + \beta_1 f_i + \beta_2 a_{pi} + \beta_3 n_i + \beta_4 f_i a_{pi} + \beta_5 f_i n_i + \beta_6 n_i a_{pi} + \beta_7 f_i n_i a_{pi} + \varepsilon_i$$

Η παραπάνω εξίσωση αναπαριστά ένα εκτενέστερο μοντέλο του προηγούμενου. Στον πίνακα 4.2, τα μοντέλα B, C και D δείχνουν το φαινόμενο αλληλεπίδρασης των τριών μεταβλητών για κάθε ένα ζεύγος τη φορά ενώ



το μοντέλο E για όλα τα ζεύγη αλληλεπίδρασης μαζί. Καθώς το μοντέλο E μπορεί να εξηγήσει μεγαλύτερη διακύμανση δεδομένων από το μοντέλο D, το μοντέλο E μπορεί να επιλεγεί ως το προτιμώμενο μοντέλο. Η παρακάτω εξίσωση λοιπόν εξηγεί το 95% της διακύμανσης των δεδομένων, αλλά αποτελεί πολυπλοκότερο μοντέλο.

$$Ra_i = -0.223 + 9.886f_i + 0.414a_{pi} - 1.93 \cdot 10^{-5} n_i - 5.91f_i a_{pi} - 5.02 \cdot 10^{-5} f_i n_i + -0.00188n_i a_{pi} + 0.00335f_i n_i a_{pi} + \varepsilon_i$$

	Βασικό Μοντέλο	Μοντέλο Αλληλεπίδρασης			
	A	B	C	D	E (βέλτιστο)
Πρόωση ( $\beta_1$ )	9.455	9.127	7.786	9.345	9.886
Βάθος κοπής ( $\beta_2$ )	0.0539	-0.0452	0.0485	-0.271	0.414
Ταχύτητα κοπής ( $\beta_3$ )	$5.61 \times 10^{-5}$	$5.56 \times 10^{-5}$	$-0.81 \times 10^{-5}$	$-0.98 \times 10^{-5}$	$-0.19 \times 10^{-5}$
Πρόωση x Βάθος κοπής ( $\beta_4$ )		0.892			-5.91
Πρόωση x Ταχύτητα κοπής ( $\beta_5$ )			0.00116		$-5 \times 10^{-5}$
Βάθος x Ταχύτητα κοπής ( $\beta_6$ )				0.000223	-0.00019
Πρόωση x βάθος x Ταχύτητα κοπής ( $\beta_7$ )					0.00335
Σταθερά	-0.279	-0.242	-0.0849	-0.164	-0.223
R <sup>2</sup>	0.925	0.924	0.928	0.947	9.95

**Πίνακας 4.2** Μοντέλα Πολλαπλής Παλινδρόμησης [1].

#### Παλινδρόμηση με Τυχαία Δάση (RFR)

Το μοντέλο RFR προσπαθεί να μειώσει τη σχέση μεταξύ συγκεντρωμένων ποσοτήτων σχεδιάζοντας ένα υποσύνολο των μεταβλητών τυχαία. Χρειάζεται δεδομένα εισόδου (τις τρεις παραμέτρους: πρόωση, βάθος κοπής και ταχύτητα κοπής και την μεταβλητή απόκρισης που είναι η επιφανειακή τραχύτητα), έναν αριθμό δέντρων (t) και έναν αριθμό μεταβλητών που θα χρησιμοποιηθούν σε κάθε "μοίρασμα" (m). Ουσιαστικά, το μοντέλο RFR προσπαθεί να προβλέψει τη μεταβλητή εξόδου (επιφανειακή τραχύτητα) από μια ομάδα πιθανών μεταβλητών πρόβλεψης (πρόωση, βάθος κοπής και ταχύτητα κοπής). Εάν μια μεταβλητή πρόβλεψης είναι σημαντική στο να κάνει την πρόβλεψη ακριβή, τότε δίνοντάς της τυχαίες τιμές, θα πρέπει να είμαστε ικανοί να αντιληφθούμε το πόσο πολύ μια πρόβλεψη μπορεί να επηρεαστεί, συγκριτικά με μια μεταβλητή που συνεισφέρει λίγο. Ο βαθμός σημαντικότητας μεταβλητής προσπαθεί να αποδόσει αυτό το φαινόμενο.

Το μοντέλο έτρεξε για t=(300,500,1000) και m=(1,2,3) και έδωσε τα καλύτερα αποτελέσματα για t=500 και m=3. Οι βαθμοί σημαντικότητας μετράνε πόσο πιο πολύ "επιδραστική" παρά "τυχαία" είναι μια συγκεκριμένη μεταβλητή στη επιτυχία πρόβλεψης της μεταβλητής εξόδου (επιφανειακή τραχύτητα). Η καλύτερη προσέγγιση δείχνει ότι η πρόωση είναι ο καλύτερος παράγοντας πρόβλεψης της επιφανειακής τραχύτητας και ακολουθούν το βάθος και η ταχύτητα κοπής [1].

#### Μοντέλο quantile παλινδρόμησης

Ενώ η πολλαπλή παλινδρόμηση παρέχει μια σύνοψη για τις μέσες τιμές των κατανομών που αντιστοιχούν στο σετ των παλινδρομητών, η quantile παλινδρόμηση βοηθά να υπολογιστούν διάφορες διαφορετικές καμπύλες παλινδρόμησης και έτσι παρέχει μια ολοκληρωμένη εικόνα των δεδομένων. Η εξίσωση θα μετατραπεί για το τ-οστό quantile στην παρακάτω σχέση:

$$Ra_i = a^\tau + \beta_1 f_i^\tau + \beta_2 d_i^\tau + \beta_3 n_i^\tau + \varepsilon_i^\tau$$

Επομένως το μοντέλο quantile παλινδρόμησης μπορεί να θεωρηθεί ως μια επέκταση του μοντέλου πολλαπλής παλινδρόμησης. Τέλος, να αναφερθεί ότι το μοντέλο quantile παλινδρόμησης δείχνει ότι η επιρροή της πρόωσης στην επιφανειακή τραχύτητα είναι μικρότερη για μικρότερες τιμές επιφανειακής τραχύτητας και μεγαλύτερη για μεγάλες τιμές επιφανειακής τραχύτητας.

Το μοντέλο quantile παλινδρόμησης δε θα συγκριθεί με τα υπόλοιπα μοντέλα καθώς η τεχνική αυτή χρησιμοποιείται για να γίνει κατανοητό πως η επιρροή των μεταβλητών πρόβλεψης είναι διαφορετική σε διαφορετικά quantiles της επιφανειακής τραχύτητας. [1]

#### Παρατηρήσεις-Συμπεράσματα

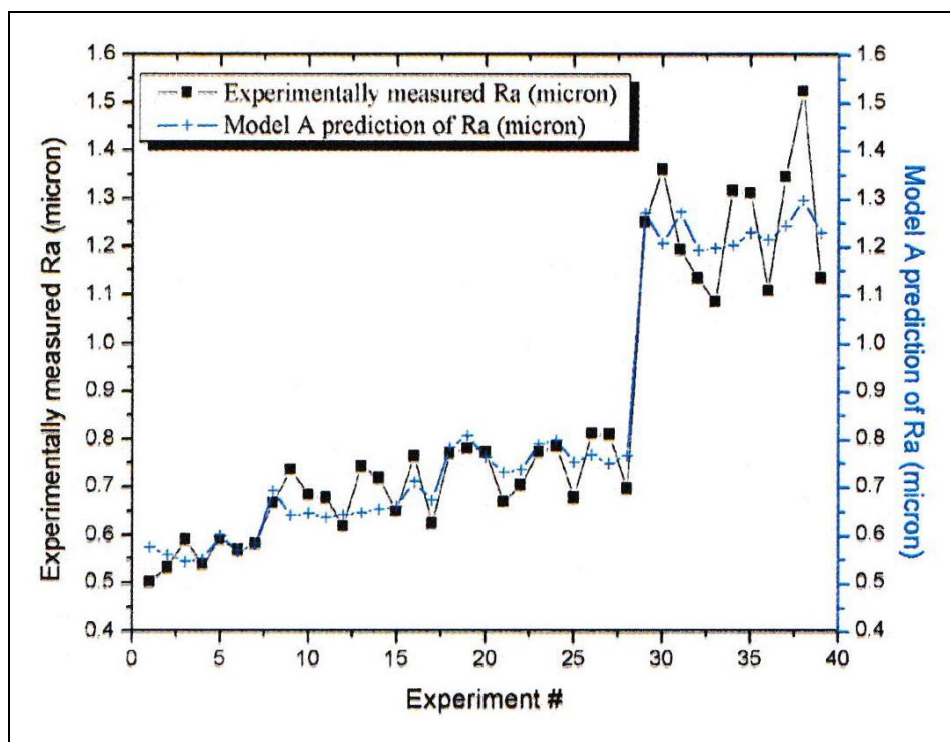
Από τα διαγράμματα 4.1 - 4.3 φαίνεται ότι παρά το γεγονός ότι τα μοντέλα A, E και των τυχαίων δασών είναι καλά στην πρόβλεψη της επιφανειακής τραχύτητας, ήταν πιο ακριβή μόνο όταν η επιφανειακή τραχύτητα βρισκόταν κάτω από την τιμή 1μm. Ενώ η επιφανειακή τραχύτητα χειροτερεύει πάνω από το 1μm και για τα 3 μοντέλα, το μοντέλο E γίνεται πιο ακριβές από το μοντέλο A. **Το μοντέλο τυχαίων δασών δείχνει συνολικά μεγαλύτερη ακρίβεια σε σχέση με τα μοντέλα A και E.** Παρακάτω παρατίθεται ένας πίνακας με τις αποκλίσεις των 3 μοντέλων:

	Μοντέλο A	Μοντέλο E	Μοντέλο Τυχαίων Δασών
Απόλυτη απόκλιση των πειραματικών δεδομένων με τις προβλεπόμενες τιμές συνολικά	0.074	0.0565	0.0465
Απόλυτη απόκλιση των πειραματικών δεδομένων με τις προβλεπόμενες τιμές για επιφανειακή τραχύτητα κάτω από 1μm	0.0479	0.0447	0.0298

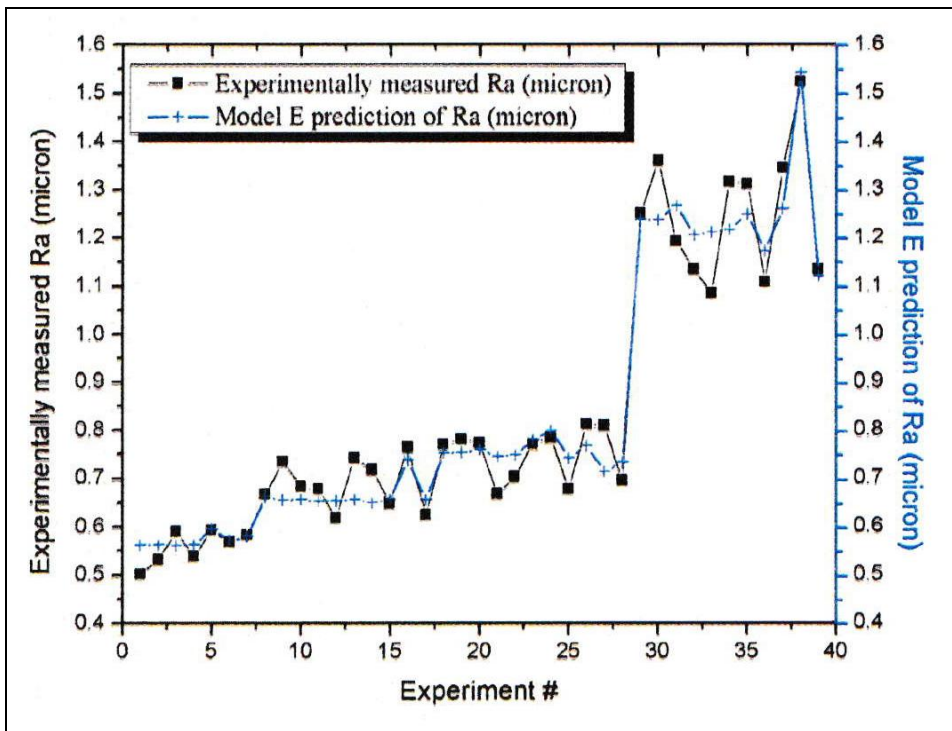
**Πίνακας 4.3** Αποκλίσεις των μοντέλων A, E και του μοντέλου τυχαίων δασών [1].

Να σημειωθεί ότι ο αριθμός πειραματικών δεδομένων είναι μικρός (39 συνολικά) και ότι ο χρόνος προσομοίωσης του μοντέλου πολλαπλών παλινδρομήσεων χρειάζεται λιγότερο χρόνο σε σχέση με το μοντέλο παλινδρόμησης με τυχαία δαση [1].

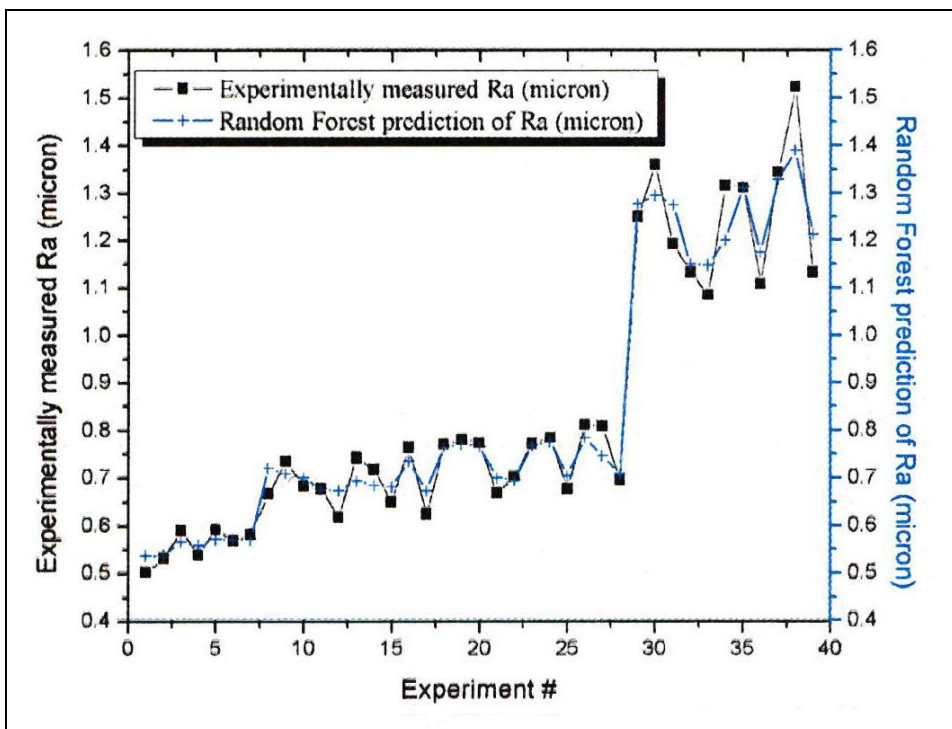
Στη συνέχεια παρατίθενται τα διαγράμματα 4.1 - 4.3 των 3 μοντέλων που αναφέρονται στα τελικά συμπεράσματα δίνοντας μεγαλύτερη προσοχή στα αποτελέσματα του μοντέλου παλινδρόμησης με τυχαία δάση το οποίο έχει το μικρότερο σφάλμα και θα συγκριθεί με τα ΤΝΔ που αναπτύξαμε εμείς.



**Διάγραμμα 4.1** Πειραματικές και προβλεπόμενες από το μοντέλο A τιμές της επιφανειακής τραχύτητας [1].



Διάγραμμα 4.2 Πειραματικές και προβλεπόμενες από το μοντέλο E τιμές της επιφανειακής τραχύτητας [1].



Διάγραμμα 4.3 Πειραματικές και προβλεπόμενες από το μοντέλο RFR τιμές της επιφανειακής τραχύτητας [1].

### 4.3 Τεχνητά νευρωνικά δίκτυα

Στην παρούσα εργασία προσπαθήσαμε να αναπτύξουμε ένα νευρωνικό δίκτυο πρόσθιας τροφοδότησης (MLP) στην MATLAB το οποίο θα μπορούσε να εκπαιδευτεί κατάλληλα ώστε να προβλέπει την επιφανειακή τραχύτητα καλύτερα από τη μέθοδο RFR που αναφέραμε παραπάνω. Ο λόγος που οι δοκιμές έγιναν σε δίκτυα πρόσθιας τροφοδότησης είναι πως η βιβλιογραφία αναφέρεται κυρίως σε αυτή την αρχιτεκτονική με ικανοποιητικά αποτελέσματα συγκριτικά με άλλες αρχιτεκτονικές. Για να καταλήξουμε σε ένα τέτοιο ΤΝΔ δοκιμάσαμε πολλούς διαφορετικούς συνδυασμούς αριθμών νευρώνων για δίκτυα με δυο κρυφά επίπεδα. Έχει αναφερθεί πως τα δύο κρυφά επίπεδα αυξάνουν την πολυπλοκότητα του δικτύου σε αντίστοιχα προβλήματα πρόβλεψης της επιφανειακής τραχύτητας και δημιουργείται πρόβλημα στη γενίκευση του δικτύου αλλά επίσης καθιστούν και χρονοβόρα την εκπαίδευσή του [14].

Καταρχάς να τονίσουμε εδώ ότι δεν μας ενδιέφερε τόσο ο χρόνος εκπαίδευσης καθώς τα σημερινά υπολογιστικά συστήματα έχουν μεγάλη επεξεργαστική ισχύ και ταχύτητα η οποία χρόνο με το χρόνο μεγαλώνει σε βαθμό που καθιστούν αυτόν τον παράγοντα λιγότερο σημαντικό. Η παρούσα εργασία δε θέλαμε να περιορίζεται από έναν τέτοιο παράγοντα που μελλοντικά θα την καθιστούσε παρωχημένη.

Όσον αφορά τη γενίκευση των δικτύων, η χρήση δυο κρυφών επιπέδων μπορεί να μας οδηγήσει κυρίως σε κίνδυνο υπερεκπαίδευσης, δηλαδή το δίκτυο να φθάνει σε σχεδόν μηδενικό σφάλμα εκπαίδευσης καθώς η συνάρτηση που προσεγγίζει γίνεται όλο και περισσότερο πολύπλοκη για να περνά από όλα τα σημεία των παραδειγμάτων εκπαίδευσης, με μεγάλη πιθανότητα τα σημεία ελέγχου να μην είναι σημεία της προσεγγιστικής αυτής συνάρτησης, δηλαδή το δίκτυο να μη γενικεύει καλά. Ωστόσο, αυτό δεν σημαίνει ότι τα δίκτυα θα υπερεκπαιδευτούν οπωσδήποτε (μάλιστα υπάρχει ένα κριτήριο κατά την εκπαίδευση που βοηθά σε αυτή τη κατεύθυνση), αλλά επίσης κάποια πιθανώς να κάνουν ανεπαρκή προσέγγιση της συνάρτησης που ψάχνουμε [5,14].

Για το θέμα της γενίκευσης, στην παρούσα εργασία ακολουθήσαμε την παρακάτω διαδικασία:

Για κάθε αρχιτεκτονική, όπου εκπαιδεύσαμε 5000 ΤΝΔ, φροντίσαμε να φιλτράρουμε τα εκπαιδευμένα αυτά δίκτυα εντοπίζοντας και αποθηκεύοντας εκείνο το οποίο είχε το μικρότερο σφάλμα ελέγχου (δηλαδή εκείνο που θεωρητικά γενικεύει καλύτερα). Ο λόγος που επιλέξαμε τον αριθμό 5000 είναι επειδή παρατηρήσαμε στατιστικά ότι μετά από 5000 νέες εκπαιδεύσεις προκύπτει υπερβολικά σπάνια κάποιο βέλτιστο (ως προς το σφάλμα ελέγχου) ΤΝΔ.

Έπειτα, για την επιλογή της βέλτιστης αρχιτεκτονικής τα αποτελέσματα προσεγγίστηκαν με δύο τρόπους-τεχνικές. Στην πρώτη τεχνική αρχικά πραγματοποιήσαμε διερεύνηση βασισμένη στη σύγκριση των μέσων τετραγωνικών σφαλμάτων (δίνοντας μεγαλύτερη βαρύτητα στο μέσο τετραγωνικό σφάλμα ελέγχου) μεταξύ των δομών αυτών και έπειτα επιλέξαμε ως βέλτιστες τις δομές με τα λιγότερα δεδομένα που είχαν σφάλμα >10% της ζητούμενης επιφανειακής τραχύτητας. Κατά τη δεύτερη τεχνική εντοπίσαμε πρώτα τις δομές με τα λιγότερα δεδομένα που είχαν σφάλμα >10% της ζητούμενης επιφανειακής τραχύτητας και στη συνέχεια προχωρήσαμε σε περαιτέρω διερεύνηση βασισμένη στη σύγκριση των μέσων τετραγωνικών σφαλμάτων (δίνοντας μεγαλύτερη βαρύτητα στο μέσο τετραγωνικό σφάλμα ελέγχου) μεταξύ των δομών αυτών.

#### Προσοχή:

Η παραπάνω διαδικασία για κάθε αρχιτεκτονική γίνεται μετά τη διαδικασία της εκπαίδευσης των ΤΝΔ στην οποία πραγματοποιείται μεν έλεγχος του σφάλματος στο validation set (όπως θα δούμε παρακάτω) αλλά ο έλεγχος αυτός γίνεται για την ρύθμιση των τιμών των βαρών και των πολώσεων του δικτύου ώστε να βελτιστοποιηθεί η απόδοση του εκπαιδευόμενου ακόμα ΤΝΔ [14]. Το κριτήριο σφάλματος του validation set ουσιαστικά προσπαθεί να αποτρέψει την υπερεκπαίδευση του εκπαιδευόμενου δικτύου (άρα να οδηγήσει προς καλή γενίκευση) αλλά αυτό δε σημαίνει ότι με το κριτήριο αυτό για κάθε αρχιτεκτονική όλα τα 5000 δίκτυα έχουν το ίδιο σφάλμα ελέγχου ή ότι γενικεύουν καλά. Επομένως υπάρχει λόγος που ακολουθούμε την παραπάνω διαδικασία.

Όσον αφορά την διαδικασία της εκπαίδευσης, η προεπιλεγμένη συνάρτηση απόδοσης των δικτύων πρόσθιας τροφοδότησης στη MATLAB είναι το μέσο τετραγωνικό σφάλμα mse μεταξύ των εξόδων και των στόχων του δικτύου και ορίζεται ως εξής:

$$F = mse = \frac{1}{N} \sum_{i=1}^N (e_i)^2 = \frac{1}{N} \sum_{i=1}^N (\text{επιθυμητή έξοδος}_i - \text{έξοδος συστήματος}_i)^2$$

όπου N είναι το πλήθος των δεδομένων.

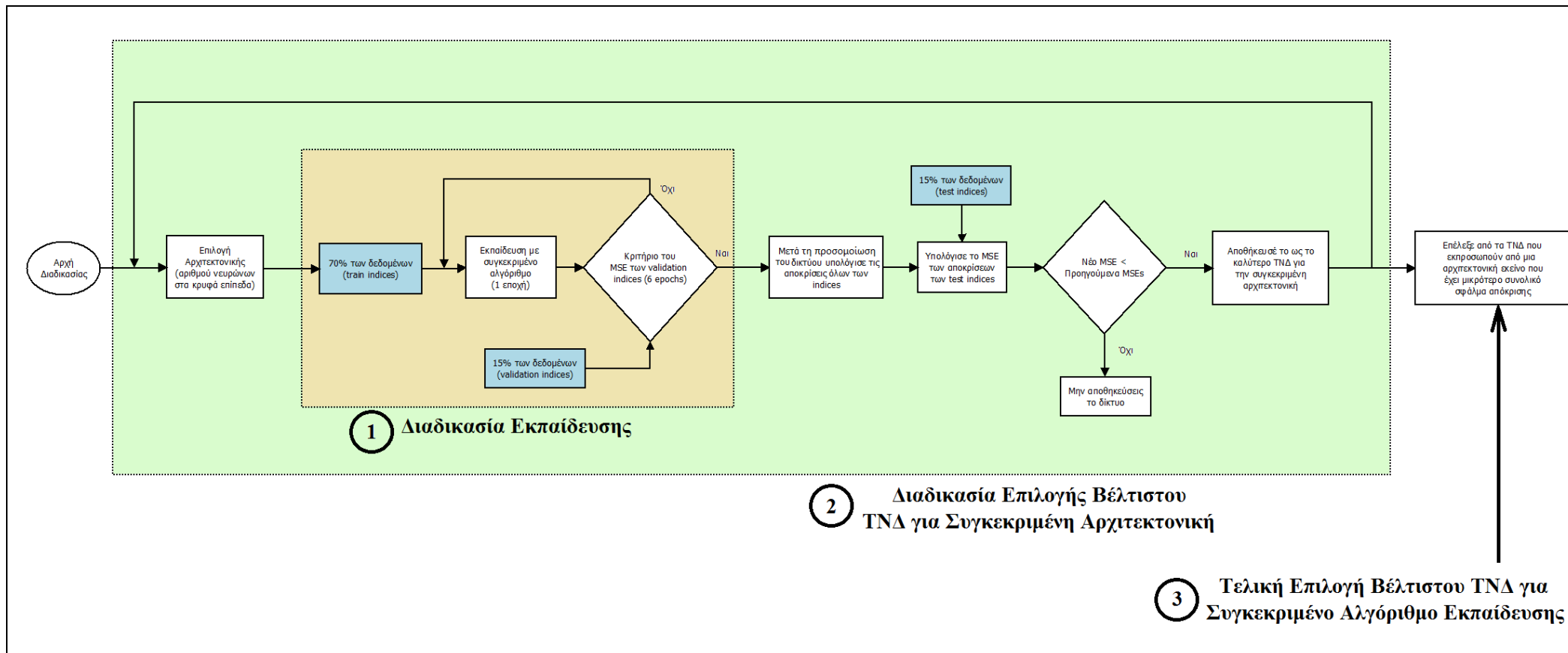
Όταν εκπαιδύουμε τα δίκτυα πρόσθιας τροφοδότησης πολλών επιπέδων, μια γενική τεχνική είναι να χωρίζουμε τα δεδομένα σε τρία υποσύνολα:

- Το πρώτο υποσύνολο είναι το **εκπαίδευσης (training set)**, το οποίο χρησιμοποιείται για τον υπολογισμό της κλίσης και την ανανέωση των βαρών και των πολώσεων του δικτύου.
- Το δεύτερο υποσύνολο είναι το επικύρωσης (**validation set**). Το σφάλμα στο validation set παρακολουθείται κατά την διαδικασία της εκπαίδευσης ως έλεγχος για να μην οδηγηθεί το δίκτυο σε υπερεκπαίδευση. Το σφάλμα αυτό κανονικά μειώνεται κατά την αρχή της εκπαίδευσης όπως και το σφάλμα του training set. Ωστόσο, όταν το δίκτυο ξεκινά να υπερμοντελοποιεί (overfit) τα δεδομένα, το σφάλμα στο validation set συνήθως αρχίζει να αυξάνεται, οπότε τα βάρη και οι πολώσεις του δικτύου αποθηκεύονται στις τιμές που είχαν πριν από τις τελευταίες 6 εποχές όπου σφάλμα του validation set άρχισε να μεγαλώνει (κριτήριο του validation set) [14,16].
- Το τρίτο υποσύνολο είναι το υποσύνολο ελέγχου (**test set**). Το σφάλμα του **test set** δεν χρησιμοποιείται κατά την εκπαίδευση. Είναι ωστόσο χρήσιμο να φαίνεται στα διαγράμματα μαζί με τα άλλα δύο σφάλματα κατά την διαδικασία εκπαίδευσης.

Αυτό ακολουθήθηκε και στην παρούσα εργασία επιλέγοντας έναν διαχωρισμό 70% / 15% / 15% αντίστοιχα όπως προτείνεται τόσο στη βιβλιογραφία όσο και από το ίδιο το MATLAB [14].

Το κριτήριο του σφάλματος του validation set έχει το εξής μειονέκτημα: Επειδή σταματά την εκπαίδευση έπειτα από 6 εποχές στις οποίες το σφάλμα του validation set δεν μειώθηκε, κάποια δίκτυα δεν σημαίνει ότι οπωσδήποτε θα οδηγούνταν σε υπερεκπαίδευση. Ουσιαστικά, εάν εντός 6 εποχών φαίνεται να οδηγούμαστε σε υπερεκπαίδευση, η εκπαίδευση του δικτύου σταματά με κίνδυνο να έχουμε δίκτυα που υπέστησαν ανεπαρκή εκπαίδευση. Ίσως ένας τρόπος αποφυγής αυτού είναι να αυξήσουμε τις εποχές από 6 σε π.χ.10 (λαμβάνοντας υπόψιν βέβαια και τον επιπλέον υπολογιστικό χρόνο), καθώς η παράληψη του κριτηρίου δεν συνιστάται.

Για καλύτερη κατανόηση της όλης διαδικασίας επιλογής του βέλτιστου ΤΝΔ υπάρχει αντίστοιχο διάγραμμα ροής στο σχήμα 4.1.



Σχήμα 4.1 Διάγραμμα Ροής της επιλογής ΤΝΔ.

Όλα τα παραπάνω επαναλήφθηκαν για τρεις διαφορετικές μεθόδους-αλγόριθμους εκπαίδευσης:

1. Τον προσαρμοστικό αλγόριθμο ανάστροφης διάδοσης της πιο απότομης κατάβασης (steepest descent) με χρήση όρου ορμής. Η μέθοδος της πλέον απότομης κατάβασης έχει αναφερθεί αναλυτικά στο θεωρητικό μέρος της εργασίας και αποτελεί μια από τις βασικότερες μεθόδους εκπαίδευσης ΤΝΔ. Στο MATLAB αναφέρεται ως "**trainGDX**".
2. Τον αλγόριθμο ανάστροφης διάδοσης των Levenberg-Marquardt. Γενικά, σε προβλήματα προσέγγισης συνάρτησης, για δίκτυα μέτρια ή μεγάλα σε μέγεθος (αριθμό βαρών) ο αλγόριθμος Levenberg-Marquardt έχει τη ταχύτερη σύγκλιση. Σε πολλές περιπτώσεις, ο αλγόριθμος αυτός είναι ικανός να δώσει τα μικρότερα μέσα τετραγωνικά σφάλματα σε σύγκριση με άλλους αλγορίθμους. Στο MATLAB αναφέρεται ως "**trainLM**".
3. Τον αλγόριθμο συζυγών κλίσεων scaled conjugate gradient. Οι αλγόριθμοι συζυγών κλίσεων (conjugate gradient algorithms) αποτελούν μια οικογένεια αλγορίθμων, οι οποίοι σε κάθε βήμα εκτελούν τοπική αναζήτηση προσπαθώντας να βρουν την βέλτιστη τιμή αλλά και την βέλτιστη κατεύθυνση μεταβολής των βαρών. Συνήθως πετυχαίνουν πολύ καλά αποτελέσματα, επιταχύνοντας κατά 10-100 φορές τη διαδικασία της εκπαίδευσης. Ο συγκεκριμένος αλγόριθμος που χρησιμοποιήσαμε, είναι ένας από τους αλγορίθμους που λειτουργούν με αναλυτικό υπολογισμό δευτέρων παραγώγων και αναφέρεται στο MATLAB ως "**trainSCG**". Άλλοι αλγόριθμοι της ίδιας οικογένειας λειτουργούν με υπολογισμό πρώτων παραγώγων και αριθμητική προσέγγιση δευτέρων παραγώγων [14,15].

Όσον αφορά την παραμετροποίηση των αλγορίθμων χρησιμοποιήσαμε τις προεπιλεγμένες τιμές καθώς αποτελούν δοκιμασμένες τιμές και στα πλαίσια της παρούσας εργασίας θα χρειαζόμασταν πολλαπλάσιο χρόνο για αντίστοιχες δοκιμές. Για όλους τους αλγορίθμους ισχύουν κάποιες κοινές παράμετροι [14,15]:

- Show Training Window Feedback      showWindow: true
- Show Command Line Feedback    showCommandLine: false
- Command Line Frequency              show: 25
- Maximum Epochs                      epochs: 1000
- Maximum Training Time              time: Inf
- Performance Goal                      goal: 0
- Maximum Validation Checks        max\_fail: 6

Ενώ παρακάτω παρατίθενται για κάθε αλγόριθμο ξεχωριστά οι τιμές των παραμέτρων του:

1. Αλγόριθμος ανάστροφης διάδοσης της πιο απότομης κατάβασης (traingdx)

- Minimum Performance Gradient      min\_grad: 1e-05
- Learning Rate                              lr: 0.01
- Learning Rate Increase                lr\_inc: 1.05
- Learning Rate                              lr\_dec: 0.7
- Maximum Performance Increase      max\_perf\_inc: 1.04
- Momentum Constant                    mc: 0.9

2. Αλγόριθμος ανάστροφης διάδοσης των Levenberg-Marquardt (trainlm)

- Minimum Performance Gradient      min\_grad: 1e-07
- Maximum Validation Checks          max\_fail: 6
- Mu    mu: 0.001
- Mu Decrease Ratio                      mu\_dec: 0.1
- Mu Increase Ratio                      mu\_inc: 10
- Maximum mu                              mu\_max: 10000000000

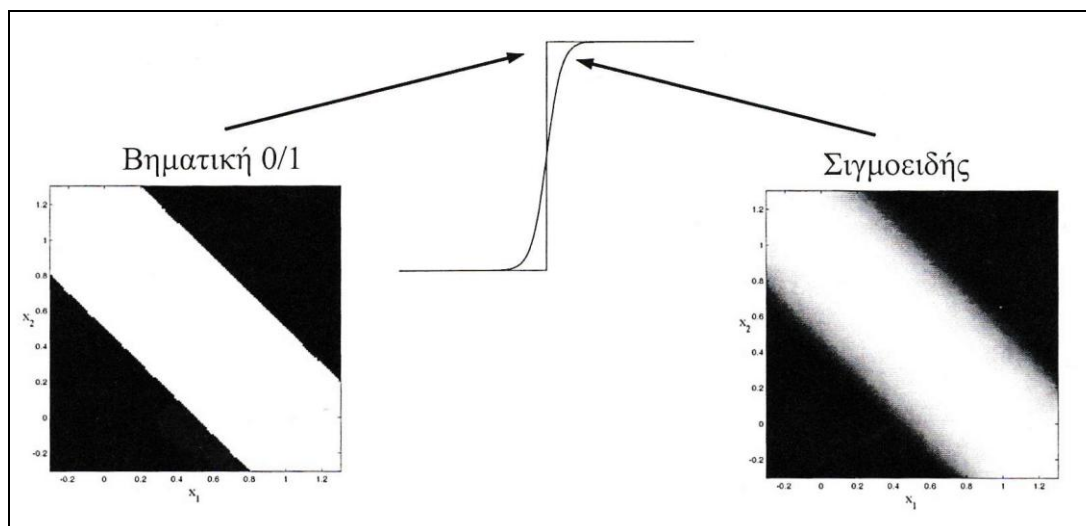
### 3. Αλγόριθμος συζυγών κλίσεων scaled conjugate gradient (trainscg)

- Minimum Performance Gradient      min\_grad: 1e-06
- Sigma      sigma: 5e-05
- Lambda      lambda: 5e-07

Η εκπαίδευση ενός νευρωνικού δικτύου σταματά σε δύο περιπτώσεις:

- Όταν η εκπαίδευση φθάσει στον μέγιστο αριθμό εποχών (έχει οριστεί στις 1000 εποχές)
- Όταν ενεργοποιηθεί το κριτήριο σφάλματος του validation set

Να σημειωθεί εδώ ότι για τα παραδείγματα εκπαίδευσης και ελέγχου του δικτύου εφαρμόστηκε η τεχνική της κανονικοποίησης δεδομένων. Πιο συγκεκριμένα, επειδή οι τρεις διαφορετικές εισόδοι και η έξοδος αποτελούν τέσσερις διαφορετικές παραμέτρους του προβλήματος, το εύρος καθώς και η φυσική έννοια των τιμών είναι διαφορετικά (π.χ. η πρόωση κυμαίνεται στο εύρος: 0.08-0.15 mm/rev, ενώ η ταχύτητα κοπής σε ένα εύρος: 858-2681 rpm). Παρότι ένα ΤΝΔ δεν ενδιαφέρεται για τα είδη (φυσικές έννοιες) των εισόδων-εξόδων, όσον αφορά στο εύρος των τιμών (ως αριθμοί), η εκπαίδευση επηρεάζεται ως προς την ταχύτητα αλλά και τη σύγκλιση. Αυτό συμβαίνει διότι η συνάρτηση ενεργοποίησης δίνει εξόδους σε ένα συγκεκριμένο εύρος [0,1]. Για παράδειγμα, έξω από μια σχετικά μικρή περιοχή γύρω του μηδέν, μια σιγμοειδής συνάρτηση φθάνει σε κορεσμό δίνοντας μόνο τιμές 0 ή 1 και ο νευρώνας συμπεριφέρεται ως ένα perceptron. Αυτό θα επηρεάσει αρνητικά την γενίκευση καθώς θα παραβλεφθούν οι ενδιάμεσες τιμές (δεν υπάρχουν "αποχρώσεις του γκρι" βλ.σχήμα 4.2). Επίσης, κατά την εκκίνηση της διαδικασίας βελτιστοποίησης των βαρών-πολώσεων, πραγματοποιείται υπολογισμός του σφάλματος ως διαφορά μεταξύ της εξόδου του δικτύου και της επιθυμητής εξόδου. Οι τιμές αυτών των δύο σετ πρέπει να κυμαίνονται στο ίδιο εύρος. Τέλος, επειδή τα μεγέθη των τιμών των δεδομένων εισόδου και εξόδου είναι διαφορετικά, πιθανότατα η εκπαίδευση να δώσει μεγαλύτερη βαρύτητα στα δεδομένα με τις μεγαλύτερες τιμές. [8]



**Σχήμα 4.2** Αριστερά: Συνάρτηση κατωφλίου. Δεξιά: Σιγμοειδής συνάρτηση [8].

Για αυτό το λόγο, τα δεδομένα κανονικοποιήθηκαν σε ένα εύρος μεταξύ 0–1 διαιρώντας, για κάθε παράμετρο, όλες τις τιμές με την μέγιστη τιμή της παραμέτρου. Τα τελικά δεδομένα που χρησιμοποιήθηκαν για την εκπαίδευση και τον έλεγχο των ΤΝΔ δίνονται στον πίνακα 4.3 στον οποίο φαίνονται και τα τρία υποσύνολα (train/validation/test) στα οποία χωρίστηκαν όπως αναφέραμε παραπάνω.



Πείραμα #i	Πρόωση (mm/rev)	Βάθος κοπής (mm)	Ταχύτητα κοπής (rpm)	Επιφανειακή τραχύτητα ( $\mu\text{m}$ )	Train / Validation / Test
1	0,5333	0,1069	0,5997	0,3296	Train
2	0,5333	0,1122	0,4662	0,3493	Train
3	0,5333	0,2139	0,3200	0,3875	Train
4	0,5333	0,2139	0,3599	0,3539	Validation
5	0,5333	0,4834	0,6900	0,3887	Test
6	0,5333	0,5796	0,3998	0,3738	Train
7	0,5333	1	0,3998	0,3822	Train
8	0,6	0,0887	0,8000	0,4379	Train
9	0,6	0,1336	0,3729	0,4826	Train
10	0,6	0,1540	0,3998	0,4484	Validation
11	0,6	0,2139	0,3200	0,4449	Test
12	0,6	0,2139	0,3599	0,4057	Train
13	0,6	0,2139	0,3998	0,4871	Train
14	0,6	0,5796	0,3599	0,4714	Train
15	0,6	0,5796	0,3998	0,4267	Validation
16	0,6	0,8053	0,7646	0,5016	Test
17	0,6	1	0,3998	0,4103	Train
18	0,6666	0,0481	0,8000	0,5055	Train
19	0,6666	0,0513	1	0,5128	Train
20	0,6666	0,1422	0,5997	0,5075	Train
21	0,6666	0,2139	0,3200	0,4390	Train
22	0,6666	0,2139	0,3599	0,4615	Train
23	0,6666	0,2502	0,8000	0,5068	Validation
24	0,6666	0,3764	0,8280	0,5147	Test
25	0,6666	0,5796	0,3998	0,4444	Train
26	0,6666	0,5967	0,5221	0,5331	Train
27	0,6666	0,8064	0,3200	0,5311	Train
28	0,6666	1	0,3998	0,4573	Train
29	1	0,0203	1	0,8214	Train
30	1	0,0641	0,4800	0,8936	Validation
31	1	0,1069	1	0,7833	Test
32	1	0,2139	0,3200	0,7445	Train
33	1	0,2139	0,3599	0,7126	Train
34	1	0,2139	0,3998	0,8640	Train
35	1	0,2973	0,5997	0,8614	Train
36	1	0,5796	0,3998	0,7277	Validation
37	1	0,7026	0,5967	0,8831	Test
38	1	0,9689	0,9697	1	Train
39	1	1	0,3998	0,7443	Train

**Πίνακας 4.4** Τα κανονικοποιημένα δεδομένα (δλδ. στο διάστημα  $[0,1]$ ) της επιφανειακής τραχύτητας για διάφορους συνδιασμούς πρόωσης, βάθους κοπής και ταχύτητας κοπής που χρησιμοποιήθηκαν για εκπαίδευση και έλεγχο των ΤΝΔ.

Εδώ να σημειώσουμε ότι γενικά η διαδικασία εύρεσης του βέλτιστου ΤΝΔ έπεται από μια σειρά διαδοχικών εκπαιδεύσεων, δεν είναι ευσταθής. Κάθε φορά που πρόκειται να ξεκινήσει μια εκπαίδευση ενός νευρωνικού δικτύου, πρέπει να δοθούν κάποιες αρχικές-τυχαίες τιμές στα αρχικά βάρη καθώς και στις πολώσεις του δικτύου [14,15].

Αυτό γίνεται στο MATLAB με την εντολή "init(net)". Η εντολή "init" καλεί την "net.initFcn" για να αρχικοποιήσει τις τιμές των βαρών και των πολώσεων σύμφωνα με τις παραμέτρους "net.initParam". Συνήθως, η "net.initFcn" είναι ορισμένη στο "initlay", που αρχικοποιεί τα βάρη και τις πολώσεις κάθε επιπέδου (layer) σύμφωνα με τη συνάρτηση "net.layers{i}.initFcn" [14,15].

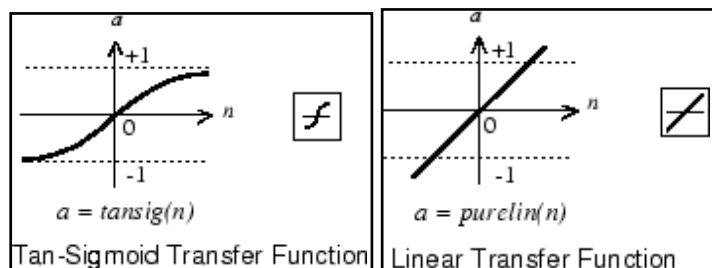
Τα δίκτυα οπισθοδιάδοσης σφάλματος (όπως τα δικά μας) έχουν την "net.layers{i}.initFcn" ορισμένη στο "initnw", και υπολογίζει τις τιμές των βαρών και των πολώσεων για κάθε επίπεδο (layer) i χρησιμοποιώντας την μέθοδο αρχικοποίησης Nguyen-Widrow [14,15].

Αυτή η τυχαία επιλογή έχει μεγάλη συνέπεια στην εκπαίδευση και το τελικό σφάλμα του δικτύου. Το ίδιο δίκτυο μπορεί να εμφανίσει επιτυχία εκπαίδευσης για μια επιλογή τυχαίων τιμών των βαρών και των πολώσεων ενώ να παρουσιάσει αποτυχία για κάποια άλλη. Για τον λόγο αυτό, για κάθε ΤΝΔ (δηλ. κάθε αρχιτεκτονική) τρέξαμε 5000 διαφορετικές εκπαιδεύσεις οι οποίες καταλήγαν σε διαφορετικές τελικές τιμές βαρών και πολώσεων. Κάθε φορά που προέκυπτε ένα βέλτιστο (ως προς το σφάλμα ελέγχου) δίκτυο, το αποθηκεύαμε στη θέση του πηγούμενου αποθηκευμένου δικτύου κ.ο.κ. Στο τέλος των 5000 εκπαιδεύσεων μιας αρχιτεκτονικής καταλήγαμε με ένα αποθηκευμένο δίκτυο για τη συγκεκριμένη αρχιτεκτονική και συνεχίζαμε ομοίως για την επόμενη κ.ο.κ.

Προφανώς υπάρχει κάποιο set αρχικών τυχαίων τιμών στα βάρη και τις πολώσεις που θα μας οδηγούσε στο τοπικό ελάχιστο (στη κοιλάδα του σφάλματος) το οποίο θα ήταν και το γενικό ελάχιστο. Ωστόσο, οι συνδυασμοί στο set αυτό είναι τόσο πολλοί που πρακτικά είναι σχεδόν αδύνατο να εξεταστούν όλοι. Ουσιαστικά με τις παραπάνω αρχικοποιήσεις, ξεκινώντας από ένα τυχαίο σημείο και ψάχνοντας το πλησιέστερο ελάχιστο, πραγματοποιούμε μια ανεπαρκή διερεύνηση ελπίζοντας να "πέσουμε" σε ένα καλό ελάχιστο (αν όχι στο γενικό ελάχιστο). Δυστυχώς, δεν υπάρχει άλλος τρόπος εξέτασης ενός νευρωνικού δικτύου.

Τα νευρωνικά δίκτυα που εξετάσαμε με τις τρεις διαφορετικές μεθόδους εκπαίδευσης είχαν ένα επίπεδο εισόδου τριών νευρώνων αφού τρεις ήταν οι είσοδοι των δεδομένων μας (πρώση, βάθος κοπής και ταχύτητα κοπής), δύο κρυμμένα επίπεδα με ποικίλο αριθμό νευρώνων (από 5 έως 20 το καθένα) και ένα επίπεδο εξόδου με ένα νευρώνα για την έξοδο του συστήματός μας που ήταν η επιφανειακή τραχύτητα.

Οι νευρώνες των κρυφών επιπέδων χρησιμοποιούσαν την σιγμοειδή συνάρτηση της υπερβολικής εφαπτομένης "tansig", ενώ το στρώμα εξόδου χρησιμοποιούσε την γραμμική συνάρτηση "purelin" καθώς για προβλήματα προσέγγισης συναρτήσεων χρησιμοποιούνται γραμμικοί νευρώνες. Οι παραπάνω συναρτήσεις φαίνονται στο σχήμα 4.3.



**Σχήμα 4.3** Αριστερά η συνάρτηση tansig των κρυφών νευρώνων και δεξιά η συνάρτηση purelin του νευρώνα εξόδου [14].

Για κάθε αλγόριθμο εκπαίδευσης, στους επόμενους πίνακες παρουσιάζονται:

- Τα αποτελέσματα της εκπαίδευσης που πήραμε για το βέλτιστο δίκτυο εκ των 5000 εκπαιδεύσεων, ανάλογα με τον αριθμό των νευρώνων στα δύο κρυφά επίπεδα. Στην πρώτη στήλη των πινάκων φαίνεται ο ακριβής αριθμός των νευρώνων των επιπέδων κάθε δικτύου. Η δεύτερη στήλη δείχνει το

μέσο τετραγωνικό σφάλμα εκπαίδευσης ενώ η τρίτη στήλη το μέσο τετραγωνικό σφάλμα των δεδομένων ελέγχου. Στην τέταρτη στήλη παρουσιάζεται ο αριθμός των εποχών εκπαίδευσης ενώ στην πεμπτη στήλη ο χρόνος που απαιτήθηκε για την εκπαίδευση σε sec. Στην τελευταία στήλη φαίνεται ο αριθμός των δεδομένων τα οποία εμφάνισαν σφάλμα μεγαλύτερο από 10% όπου έχουμε επισημάνει (με πράσινο χρώμα) τους μικρότερους αριθμούς τέτοιων δεδομένων.

- Το μέσο τετραγωνικό σφάλμα των δεδομένων ελέγχου
- Τον αριθμό εποχών που έφθασε κάθε δομή
- Τον χρόνο εκπαίδευσης κάθε δομής
- Το μέσο τετραγωνικό σφάλμα των δεδομένων εκπαίδευσης

Όσον αφορά στους 4 τελευταίους πίνακες κάθε μεθόδου:

- Κατά την οριζόντια διεύθυνση καθώς κινούμαστε προς τα δεξιά οδηγούμαστε σε δίκτυα με μεγαλύτερο αριθμό κρυφών νευρώνων στο 2ο κρυφό επίπεδο, ενώ κατά την κατακόρυφη διεύθυνση καθώς κινούμαστε προς τα κάτω οδηγούμαστε σε δίκτυα με μεγαλύτερο αριθμό κρυφών νευρώνων στο 1ο κρυφό επίπεδο.
- Οι χρωματισμοί των κελιών χρησιμοποιήθηκαν για να γίνει ευκολότερα αντιληπτή η διακύμανση των τιμών εντός κάθε πίνακα. Κινούμενοι από τη μεγαλύτερη προς τη μικρότερη τιμή, τα χρώματα αποδίδονται από το κόκκινο προς το πράσινο αντίστοιχα.

#### **Πίνακας 4.5**

Συνολικά αποτελέσματα με τη μέθοδο εκπαίδευσης της απότομης κατάβασης (trainGDX)

Αρχιτεκτονική	MSE Εκπαίδευσης	MSE Ελέγχου	Εποχές	Χρόνος (sec)	Δεδομένα με σφάλμα >10%
3-5-5-1	1,27E-03	2,62E-04	177	0,246	7
3-5-6-1	4,39E-03	3,47E-04	94	0,101	14
3-5-7-1	9,99E-04	4,30E-04	181	0,156	4
3-5-8-1	4,45E-03	4,20E-04	114	0,109	16
3-5-9-1	1,44E-03	1,57E-04	169	0,158	8
3-5-10-1	3,79E-03	3,89E-04	114	0,106	16
3-5-11-1	2,33E-03	4,54E-04	130	0,11	13
3-5-12-1	2,25E-03	4,27E-04	130	0,109	9
3-5-13-1	5,36E-03	3,49E-04	101	0,078	17
3-5-14-1	1,37E-03	2,22E-04	141	0,125	7
3-5-15-1	1,21E-02	2,92E-04	80	0,063	21
3-5-16-1	7,48E-03	1,56E-04	98	0,079	20
3-5-17-1	1,27E-03	2,94E-04	156	0,156	4
3-5-18-1	1,39E-03	4,91E-04	154	0,14	7
3-5-19-1	1,46E-03	5,88E-04	153	0,14	8
3-5-20-1	2,34E-03	3,28E-04	95	0,078	10
3-6-5-1	1,14E-03	2,05E-04	182	0,172	6
3-6-6-1	1,32E-03	1,80E-04	157	0,156	4
3-6-7-1	7,04E-03	2,15E-04	98	0,078	11
3-6-8-1	5,36E-03	4,57E-04	81	0,063	13
3-6-9-1	1,69E-03	2,44E-04	165	0,14	4
3-6-10-1	1,23E-03	1,78E-04	146	0,127	5
3-6-11-1	5,10E-03	3,51E-04	88	0,078	15
3-6-12-1	2,34E-03	4,94E-04	133	0,11	10
3-6-13-1	2,11E-03	4,17E-04	165	0,14	12

3-6-14-1	2,40E-03	4,00E-04	178	0,156	13
3-6-15-1	1,41E-03	4,22E-04	134	0,109	4
3-6-16-1	2,27E-03	3,47E-04	147	0,125	7
3-6-17-1	1,44E-03	2,57E-04	113	0,094	4
3-6-18-1	2,22E-03	6,65E-04	122	0,093	10
3-6-19-1	2,94E-03	4,40E-04	97	0,078	8
3-6-20-1	1,36E-03	5,26E-04	151	0,14	9
3-7-5-1	7,05E-03	2,64E-04	96	0,094	17
3-7-6-1	4,48E-03	2,21E-04	104	0,079	13
3-7-7-1	1,80E-03	2,05E-04	144	0,124	9
3-7-8-1	1,29E-03	4,36E-04	169	0,141	8
3-7-9-1	1,65E-03	3,47E-04	171	0,174	8
3-7-10-1	1,55E-03	4,37E-04	181	0,175	10
3-7-11-1	1,25E-03	5,51E-04	154	0,124	6
3-7-12-1	1,54E-03	4,07E-04	122	0,125	8
3-7-13-1	4,19E-03	3,42E-04	104	0,093	11
3-7-14-1	8,91E-04	1,84E-04	184	0,156	4
3-7-15-1	1,74E-03	4,27E-04	126	0,11	7
3-7-16-1	1,01E-03	2,62E-04	145	0,126	4
3-7-17-1	2,15E-03	6,01E-05	146	0,125	7
3-7-18-1	1,18E-03	3,73E-04	131	0,125	6
3-7-19-1	1,56E-03	3,41E-04	159	0,14	8
3-7-20-1	1,16E-03	3,78E-04	128	0,125	4
3-8-5-1	4,33E-03	2,29E-04	94	0,078	13
3-8-6-1	2,45E-03	2,79E-04	130	0,127	10
3-8-7-1	1,02E-03	3,36E-04	162	0,125	5
3-8-8-1	1,30E-03	2,87E-04	141	0,124	11
3-8-9-1	7,90E-04	3,92E-04	168	0,14	3
3-8-10-1	2,17E-03	4,54E-04	134	0,14	8
3-8-11-1	1,39E-03	1,70E-04	138	0,125	9
3-8-12-1	9,94E-04	3,94E-04	188	0,171	2
3-8-13-1	2,12E-03	4,91E-04	143	0,109	8
3-8-14-1	7,47E-02	2,65E-04	37	0,047	23
3-8-15-1	1,77E-03	2,44E-04	118	0,094	7
3-8-16-1	2,03E-03	3,61E-04	162	0,14	7
3-8-17-1	2,71E-03	3,84E-04	111	0,094	11
3-8-18-1	1,56E-03	2,20E-04	154	0,142	5
3-8-19-1	7,45E-03	1,88E-04	82	0,062	18
3-8-20-1	1,55E-03	1,17E-04	170	0,156	9
3-9-5-1	2,41E-03	1,88E-04	131	0,109	10
3-9-6-1	4,83E-03	3,43E-04	106	0,093	10
3-9-7-1	1,29E-03	2,87E-04	133	0,126	6
3-9-8-1	2,04E-03	2,64E-04	136	0,111	12
3-9-9-1	1,61E-03	3,86E-04	120	0,111	6
3-9-10-1	7,98E-04	3,00E-04	177	0,156	4
3-9-11-1	7,38E-04	3,15E-04	175	0,141	2
3-9-12-1	1,75E-03	2,10E-04	145	0,125	8
3-9-13-1	7,16E-03	3,39E-04	90	0,078	14

3-9-14-1	9,00E-03	4,12E-04	57	0,062	20
3-9-15-1	5,99E-04	6,29E-04	162	0,141	5
3-9-16-1	9,26E-04	3,93E-04	164	0,14	2
3-9-17-1	8,81E-04	4,13E-04	156	0,125	4
3-9-18-1	2,43E-03	4,92E-04	102	0,11	11
3-9-19-1	1,39E-02	4,31E-04	88	0,078	20
3-9-20-1	2,42E-02	5,66E-04	61	0,062	26
3-10-5-1	3,43E-03	3,05E-04	118	0,109	14
3-10-6-1	2,12E-03	3,37E-04	121	0,106	10
3-10-7-1	2,85E-03	2,97E-04	104	0,093	11
3-10-8-1	2,01E-03	2,70E-04	142	0,126	9
3-10-9-1	2,29E-03	3,51E-04	100	0,094	8
3-10-10-1	2,69E-03	3,56E-04	108	0,094	11
3-10-11-1	7,57E-04	2,62E-04	145	0,125	2
3-10-12-1	2,04E-03	4,70E-04	124	0,11	6
3-10-13-1	2,18E-03	3,32E-04	116	0,109	10
3-10-14-1	8,18E-04	5,34E-04	160	0,157	4
3-10-15-1	1,60E-03	3,80E-04	117	0,125	9
3-10-16-1	1,02E-03	5,21E-04	181	0,14	7
3-10-17-1	2,44E-03	2,78E-04	121	0,109	15
3-10-18-1	9,55E-04	7,64E-04	181	0,14	1
3-10-19-1	2,69E-03	5,51E-04	116	0,109	11
3-10-20-1	3,10E-03	3,78E-04	100	0,078	9
3-11-5-1	1,42E-03	2,81E-04	144	0,129	4
3-11-6-1	2,12E-03	2,18E-04	113	0,109	8
3-11-7-1	1,03E-03	5,11E-04	166	0,157	2
3-11-8-1	1,88E-03	4,05E-04	112	0,093	11
3-11-9-1	5,48E-03	1,55E-04	86	0,078	15
3-11-10-1	4,63E-03	4,27E-04	118	0,11	15
3-11-11-1	2,71E-03	2,95E-04	127	0,109	9
3-11-12-1	7,69E-04	6,49E-04	166	0,14	3
3-11-13-1	1,17E-03	3,69E-04	157	0,125	7
3-11-14-1	1,83E-03	3,61E-04	118	0,094	8
3-11-15-1	1,56E-03	4,32E-04	111	0,109	6
3-11-16-1	2,62E-03	2,80E-04	107	0,109	9
3-11-17-1	3,65E-03	8,46E-04	81	0,078	20
3-11-18-1	2,33E-03	2,91E-04	108	0,094	10
3-11-19-1	2,13E-03	4,44E-04	124	0,109	6
3-11-20-1	6,90E-04	4,67E-04	155	0,14	2
3-12-5-1	9,49E-04	3,72E-04	168	0,142	5
3-12-6-1	9,22E-04	3,25E-04	173	0,172	3
3-12-7-1	2,66E-03	4,03E-04	117	0,096	11
3-12-8-1	1,97E-03	2,23E-04	147	0,125	10
3-12-9-1	1,09E-03	2,12E-04	166	0,14	4
3-12-10-1	2,74E-03	3,10E-04	114	0,094	12
3-12-11-1	1,83E-03	4,39E-04	134	0,125	9
3-12-12-1	8,09E-04	1,66E-04	165	0,156	3
3-12-13-1	1,15E-03	3,08E-04	162	0,14	5

3-12-14-1	3,69E-04	3,86E-04	181	0,14	1
3-12-15-1	1,05E-03	2,97E-04	132	0,125	4
3-12-16-1	1,29E-03	3,53E-04	138	0,14	7
3-12-17-1	1,22E-03	3,69E-04	130	0,126	6
3-12-18-1	2,20E-03	3,42E-04	106	0,093	10
3-12-19-1	1,37E-03	5,09E-04	115	0,093	8
3-12-20-1	1,07E-03	4,11E-04	150	0,141	7
3-13-5-1	3,33E-03	2,15E-04	113	0,093	10
3-13-6-1	5,99E-03	1,65E-04	104	0,094	12
3-13-7-1	6,13E-04	3,55E-04	175	0,156	1
3-13-8-1	3,24E-03	1,70E-04	117	0,205	11
3-13-9-1	2,08E-03	9,25E-05	125	0,125	10
3-13-10-1	2,61E-03	2,02E-04	116	0,093	9
3-13-11-1	5,67E-03	3,03E-04	87	0,078	12
3-13-12-1	7,75E-04	3,48E-04	165	0,141	3
3-13-13-1	2,02E-03	4,70E-04	128	0,124	13
3-13-14-1	3,85E-03	4,55E-04	108	0,094	16
3-13-15-1	1,78E-03	9,39E-05	104	0,116	9
3-13-16-1	8,46E-04	3,12E-04	156	0,146	5
3-13-17-1	2,35E-03	2,70E-04	116	0,109	10
3-13-18-1	7,60E-04	2,66E-04	148	0,157	1
3-13-19-1	1,49E-03	3,15E-04	111	0,094	9
3-13-20-1	5,54E-04	4,38E-04	158	0,156	1
3-14-5-1	1,98E-03	2,18E-04	132	0,125	8
3-14-6-1	8,69E-04	2,05E-04	175	0,156	5
3-14-7-1	1,92E-03	1,94E-04	128	0,111	10
3-14-8-1	2,14E-03	3,36E-04	138	0,141	11
3-14-9-1	1,30E-03	2,02E-04	124	0,109	8
3-14-10-1	2,81E-03	2,70E-04	128	0,11	13
3-14-11-1	1,09E-03	4,17E-04	143	0,128	6
3-14-12-1	3,74E-03	3,78E-04	94	0,345	10
3-14-13-1	6,62E-04	3,74E-04	155	0,141	1
3-14-14-1	1,08E-03	3,43E-04	109	0,12	4
3-14-15-1	1,42E-03	2,58E-04	120	0,106	6
3-11-16-1	2,39E-03	3,07E-04	99	0,093	11
3-14-17-1	1,57E-03	9,61E-04	108	0,094	6
3-14-18-1	2,43E-03	3,93E-04	101	0,11	9
3-14-19-1	6,84E-03	3,69E-04	76	0,062	17
3-14-20-1	4,76E-03	9,76E-04	92	0,078	14
3-15-5-1	1,61E-03	8,60E-05	141	0,125	5
3-15-6-1	8,59E-04	2,47E-04	146	0,125	5
3-15-7-1	5,31E-03	3,24E-04	93	0,078	18
3-15-8-1	4,72E-03	3,51E-04	106	0,109	11
3-15-9-1	1,70E-03	4,04E-04	129	0,124	8
3-15-10-1	6,72E-04	1,72E-04	165	0,172	5
3-15-11-1	9,51E-04	3,19E-04	163	0,14	6
3-15-12-1	1,98E-03	3,95E-04	127	0,11	10
3-15-13-1	2,75E-03	9,89E-05	85	0,075	10

3-15-14-1	1,12E-03	3,04E-04	128	0,109	5
3-15-15-1	6,47E-04	4,50E-04	138	0,125	2
3-15-16-1	5,08E-04	2,68E-04	149	0,125	1
3-15-17-1	7,97E-04	7,69E-04	156	0,156	4
3-15-18-1	5,60E-04	8,68E-04	174	0,156	5
3-15-19-1	8,49E-04	4,36E-04	135	0,14	2
3-15-20-1	9,27E-04	4,64E-04	123	0,297	4
3-16-5-1	1,21E-03	3,65E-04	190	0,171	6
3-16-6-1	1,15E-03	2,78E-04	171	0,156	9
3-16-7-1	7,26E-03	1,92E-04	82	0,078	12
3-16-8-1	1,15E-03	1,39E-04	165	0,156	6
3-16-9-1	6,68E-03	2,79E-04	68	0,063	15
3-16-10-1	6,56E-04	2,08E-04	169	0,141	2
3-16-11-1	1,48E-03	4,97E-04	128	0,109	8
3-16-12-1	2,13E-03	4,77E-04	115	0,111	6
3-16-13-1	1,21E-03	1,50E-04	139	0,109	5
3-16-14-1	1,60E-03	4,35E-04	122	0,11	6
3-16-15-1	7,70E-04	5,69E-04	170	0,156	4
3-16-16-1	4,78E-03	2,02E-04	91	0,078	15
3-16-17-1	1,30E-02	4,37E-04	66	0,062	19
3-16-18-1	1,10E-03	5,87E-04	125	0,109	5
3-16-19-1	1,88E-03	2,75E-04	93	0,078	9
3-16-20-1	7,09E-04	4,40E-04	170	0,156	3
3-17-5-1	6,96E-04	1,50E-04	191	0,156	1
3-17-6-1	1,22E-02	2,22E-04	82	0,078	22
3-17-7-1	3,22E-03	2,91E-04	96	0,078	12
3-17-8-1	4,52E-03	3,37E-04	105	0,109	14
3-17-9-1	9,21E-04	6,45E-05	159	0,14	6
3-17-10-1	4,95E-04	2,56E-04	147	0,14	2
3-17-11-1	1,40E-03	3,20E-04	131	0,109	5
3-17-12-1	6,87E-04	5,51E-04	166	0,156	3
3-17-13-1	3,29E-03	2,24E-04	97	0,078	12
3-17-14-1	6,09E-04	4,59E-04	163	0,141	2
3-17-15-1	1,43E-03	3,69E-04	124	0,109	7
3-17-16-1	1,07E-03	4,82E-04	127	0,11	6
3-17-17-1	7,06E-04	4,57E-04	156	0,141	3
3-17-18-1	1,80E-03	4,39E-04	108	0,094	10
3-17-19-1	7,16E-04	3,64E-04	174	0,171	4
3-17-20-1	9,40E-03	4,83E-04	80	0,078	17
3-18-5-1	1,62E-03	1,13E-04	160	0,14	8
3-18-6-1	6,38E-04	2,83E-04	200	0,171	3
3-18-7-1	3,53E-03	1,94E-04	113	0,094	7
3-18-8-1	1,06E-03	2,39E-04	164	0,171	5
3-18-9-1	1,65E-03	4,53E-04	114	0,094	4
3-18-10-1	1,74E-03	4,63E-04	97	0,094	9
3-18-11-1	9,75E-04	1,63E-04	165	0,14	6
3-18-12-1	1,63E-03	5,25E-04	149	0,14	8
3-18-13-1	5,72E-04	5,30E-04	148	0,157	1

3-18-14-1	1,56E-02	2,85E-04	50	0,047	20
3-18-15-1	5,39E-04	4,51E-04	158	0,156	3
3-18-16-1	8,92E-04	5,02E-04	140	0,14	3
3-18-17-1	7,24E-04	6,58E-04	167	0,171	3
3-18-18-1	1,03E-03	1,85E-04	124	0,109	4
3-18-19-1	1,34E-03	8,00E-04	104	0,109	7
3-18-20-1	2,43E-03	7,38E-04	96	0,093	14
3-19-5-1	1,16E-03	1,11E-04	123	0,109	6
3-19-6-1	1,11E-03	4,84E-04	162	0,14	6
3-19-7-1	5,27E-03	4,17E-04	80	0,078	18
3-19-8-1	2,60E-03	3,72E-04	120	0,109	10
3-19-9-1	7,10E-04	1,65E-04	170	0,172	3
3-19-10-1	4,65E-04	2,29E-04	172	0,172	1
3-19-11-1	4,53E-03	2,60E-04	86	0,078	20
3-19-12-1	5,80E-04	2,92E-04	168	0,141	3
3-19-13-1	9,97E-04	4,15E-04	164	0,156	4
3-19-14-1	6,64E-04	2,68E-04	162	0,156	3
3-19-15-1	6,87E-04	2,57E-04	179	0,156	3
3-19-16-1	7,93E-04	4,05E-04	149	0,141	5
3-19-17-1	1,49E-03	1,00E-03	103	0,094	7
3-19-18-1	7,97E-04	7,59E-04	145	0,124	2
3-19-19-1	6,12E-03	8,63E-04	70	0,078	14
3-19-20-1	4,92E-03	3,36E-04	66	0,062	15
3-20-5-1	3,38E-03	4,41E-04	100	0,093	13
3-20-6-1	3,42E-03	1,64E-04	104	0,109	13
3-20-7-1	7,52E-04	1,01E-04	162	0,141	2
3-20-8-1	1,33E-03	3,50E-04	138	0,125	6
3-20-9-1	7,58E-03	6,53E-04	76	0,063	19
3-20-10-1	5,34E-04	3,52E-04	184	0,172	3
3-20-11-1	1,96E-03	3,61E-04	126	0,109	8
3-20-12-1	8,99E-04	4,19E-04	180	0,156	4
3-20-13-1	4,58E-04	1,75E-04	152	0,125	2
3-20-14-1	5,01E-04	3,06E-04	158	0,141	1
3-20-15-1	1,47E-03	4,19E-04	111	0,109	7
3-20-16-1	5,49E-04	1,57E-04	157	0,141	2
3-20-17-1	8,53E-04	7,96E-04	141	0,124	3
3-20-18-1	1,04E-03	8,74E-04	127	0,109	4
3-20-19-1	5,55E-04	3,73E-04	147	0,125	2
3-20-20-1	6,70E-03	5,03E-04	67	0,063	11



		Αριθμός Νευρώνων 2ου Κρυφού Επιπέδου																
		5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
mse δεδομένων ελέγχου	Αριθμός Νευρώνων 1ου Κρυφού Επιπέδου	5	2,62E-04	3,47E-04	4,30E-04	4,20E-04	1,57E-04	3,89E-04	4,54E-04	4,27E-04	3,49E-04	2,22E-04	2,92E-04	1,56E-04	2,94E-04	4,91E-04	5,88E-04	3,28E-04
		6	2,05E-04	1,80E-04	2,15E-04	4,57E-04	2,44E-04	1,78E-04	3,51E-04	4,94E-04	4,17E-04	4,00E-04	4,22E-04	3,47E-04	2,57E-04	6,65E-04	4,40E-04	5,26E-04
		7	2,64E-04	2,21E-04	2,05E-04	4,36E-04	3,47E-04	4,37E-04	5,51E-04	4,07E-04	3,42E-04	1,84E-04	4,27E-04	2,62E-04	6,01E-05	3,73E-04	3,41E-04	3,78E-04
		8	2,29E-04	2,79E-04	3,36E-04	2,87E-04	3,92E-04	4,54E-04	1,70E-04	3,94E-04	4,91E-04	2,65E-04	2,44E-04	3,61E-04	3,84E-04	2,20E-04	1,88E-04	1,17E-04
		9	1,88E-04	3,43E-04	2,87E-04	2,64E-04	3,86E-04	3,00E-04	3,15E-04	2,10E-04	3,39E-04	4,12E-04	6,29E-04	3,93E-04	4,13E-04	4,92E-04	4,31E-04	5,66E-04
		10	3,05E-04	3,37E-04	2,97E-04	2,70E-04	3,51E-04	3,56E-04	2,62E-04	4,70E-04	3,32E-04	5,34E-04	3,80E-04	5,21E-04	2,78E-04	7,64E-04	5,51E-04	3,78E-04
		11	2,81E-04	2,18E-04	5,11E-04	4,05E-04	1,55E-04	4,27E-04	2,95E-04	6,49E-04	3,69E-04	3,61E-04	4,32E-04	2,80E-04	8,46E-04	2,91E-04	4,44E-04	4,67E-04
		12	3,72E-04	3,25E-04	4,03E-04	2,23E-04	2,12E-04	3,10E-04	4,39E-04	1,66E-04	3,08E-04	3,86E-04	2,97E-04	3,53E-04	3,69E-04	3,42E-04	5,09E-04	4,11E-04
		13	2,15E-04	1,65E-04	3,55E-04	1,70E-04	9,25E-05	2,02E-04	3,03E-04	3,48E-04	4,70E-04	4,55E-04	9,39E-05	3,12E-04	2,70E-04	2,66E-04	3,15E-04	4,38E-04
		14	2,18E-04	2,05E-04	1,94E-04	3,36E-04	2,02E-04	2,70E-04	4,17E-04	3,78E-04	3,74E-04	3,43E-04	2,58E-04	3,07E-04	9,61E-04	3,93E-04	3,69E-04	9,76E-04
		15	8,60E-05	2,47E-04	3,24E-04	3,51E-04	4,04E-04	1,72E-04	3,19E-04	3,95E-04	9,89E-05	3,04E-04	4,50E-04	2,68E-04	7,69E-04	8,68E-04	4,36E-04	4,64E-04
		16	3,65E-04	2,78E-04	1,92E-04	1,39E-04	2,79E-04	2,08E-04	4,97E-04	4,77E-04	1,50E-04	4,35E-04	5,69E-04	2,02E-04	4,37E-04	5,87E-04	2,75E-04	4,40E-04
		17	1,50E-04	2,22E-04	2,91E-04	3,37E-04	6,45E-05	2,56E-04	3,20E-04	5,51E-04	2,24E-04	4,59E-04	3,69E-04	4,82E-04	4,57E-04	4,39E-04	3,64E-04	4,83E-04
		18	1,13E-04	2,83E-04	1,94E-04	2,39E-04	4,53E-04	4,63E-04	1,63E-04	5,25E-04	5,30E-04	2,85E-04	4,51E-04	5,02E-04	6,58E-04	1,85E-04	8,00E-04	7,38E-04
		19	1,11E-04	4,84E-04	4,17E-04	3,72E-04	1,65E-04	2,29E-04	2,60E-04	2,92E-04	4,15E-04	2,68E-04	2,57E-04	4,05E-04	1,00E-03	7,59E-04	8,63E-04	3,36E-04
		20	4,41E-04	1,64E-04	1,01E-04	3,50E-04	6,53E-04	3,52E-04	3,61E-04	4,19E-04	1,75E-04	3,06E-04	4,19E-04	1,57E-04	7,96E-04	8,74E-04	3,73E-04	5,03E-04

Πίνακας 4.6 Μέσο τετραγωνικό σφάλμα δεδομένων ελέγχου με την μέθοδο trainGDx.

		Αριθμός Νευρώνων 2ου Κρυφού Επιπέδου																
		5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
Αριθμός Εποχών Εκπαίδευσης	Αριθμός Νευρώνων 1ου Κρυφού Επιπέδου	5	177	94	181	114	169	114	130	130	101	141	80	98	156	154	153	95
		6	182	157	98	81	165	146	88	133	165	178	134	147	113	122	97	151
		7	96	104	144	169	171	181	154	122	104	184	126	145	146	131	159	128
		8	94	130	162	141	168	134	138	188	143	37	118	162	111	154	82	170
		9	131	106	133	136	120	177	175	145	90	57	162	164	156	102	88	61
		10	118	121	104	142	100	108	145	124	116	160	117	181	121	181	116	100
		11	144	113	166	112	86	118	127	166	157	118	111	107	81	108	124	155
		12	168	173	117	147	166	114	134	165	162	181	132	138	130	106	115	150
		13	113	104	175	117	125	116	87	165	128	108	104	156	116	148	111	158
		14	132	175	128	138	124	128	143	94	155	109	120	99	108	101	76	92
		15	141	146	93	106	129	165	163	127	85	128	138	149	156	174	135	123
		16	190	171	82	165	68	169	128	115	139	122	170	91	66	125	93	170
		17	191	82	96	105	159	147	131	166	97	163	124	127	156	108	174	80
		18	160	200	113	164	114	97	165	149	148	50	158	140	167	124	104	96
		19	123	162	80	120	170	172	86	168	164	162	179	149	103	145	70	66
		20	100	104	162	138	76	184	126	180	152	158	111	157	141	127	147	67

Πίνακας 4.7 Αριθμός εποχών που έφθασε κάθε δομή με την μέθοδο trainGDx.

Χρόνος Εκπαίδευσης σε sec	Αριθμός Νευρώνων 1ου Κρυφού Επιπέδου	Αριθμός Νευρώνων 2ου Κρυφού Επιπέδου															
		5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
		5	0,246	0,101	0,156	0,109	0,158	0,106	0,11	0,109	0,078	0,125	0,063	0,079	0,156	0,14	0,14
6	0,172	0,156	0,078	0,063	0,14	0,127	0,078	0,11	0,14	0,156	0,109	0,125	0,094	0,093	0,078	0,14	
7	0,094	0,079	0,124	0,141	0,174	0,175	0,124	0,125	0,093	0,156	0,11	0,126	0,125	0,125	0,14	0,125	
8	0,078	0,127	0,125	0,124	0,14	0,14	0,125	0,171	0,109	0,047	0,094	0,14	0,094	0,142	0,062	0,156	
9	0,109	0,093	0,126	0,111	0,111	0,156	0,141	0,125	0,078	0,062	0,141	0,14	0,125	0,11	0,078	0,062	
10	0,109	0,106	0,093	0,126	0,094	0,094	0,125	0,11	0,109	0,157	0,125	0,14	0,109	0,14	0,109	0,078	
11	0,129	0,109	0,157	0,093	0,078	0,11	0,109	0,14	0,125	0,094	0,109	0,109	0,078	0,094	0,109	0,14	
12	0,142	0,172	0,096	0,125	0,14	0,094	0,125	0,156	0,14	0,14	0,125	0,14	0,126	0,093	0,093	0,141	
13	0,093	0,094	0,156	0,205	0,125	0,093	0,078	0,141	0,124	0,094	0,116	0,146	0,109	0,157	0,094	0,156	
14	0,125	0,156	0,111	0,141	0,109	0,11	0,128	0,345	0,141	0,12	0,106	0,093	0,094	0,11	0,062	0,078	
15	0,125	0,125	0,078	0,109	0,124	0,172	0,14	0,11	0,075	0,109	0,125	0,125	0,156	0,156	0,14	0,297	
16	0,171	0,156	0,078	0,156	0,063	0,141	0,109	0,111	0,109	0,11	0,156	0,078	0,062	0,109	0,078	0,156	
17	0,156	0,078	0,078	0,109	0,14	0,14	0,109	0,156	0,078	0,141	0,109	0,11	0,141	0,094	0,171	0,078	
18	0,14	0,171	0,094	0,171	0,094	0,094	0,14	0,14	0,157	0,047	0,156	0,14	0,171	0,109	0,109	0,093	
19	0,109	0,14	0,078	0,109	0,172	0,172	0,078	0,141	0,156	0,156	0,156	0,141	0,094	0,124	0,078	0,062	
20	0,093	0,109	0,141	0,125	0,063	0,172	0,109	0,156	0,125	0,141	0,109	0,141	0,124	0,109	0,125	0,063	

Πίνακας 4.8 Χρόνος εκπαίδευσης κάθε δομής με την μέθοδο trainGDX.

		Αριθμός Νευρώνων 2ου Κρυφού Επιπέδου																
		5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
mse δεδομένων εκπαίδευσης	Αριθμός Νευρώνων 1ου Κρυφού Επιπέδου	5	1,27E-03	4,39E-03	9,99E-04	4,45E-03	1,44E-03	3,79E-03	2,33E-03	2,25E-03	5,36E-03	1,37E-03	1,21E-02	7,48E-03	1,27E-03	1,39E-03	1,46E-03	2,34E-03
		6	1,14E-03	1,32E-03	7,04E-03	5,36E-03	1,69E-03	1,23E-03	5,10E-03	2,34E-03	2,11E-03	2,40E-03	1,41E-03	2,27E-03	1,44E-03	2,22E-03	2,94E-03	1,36E-03
		7	7,05E-03	4,48E-03	1,80E-03	1,29E-03	1,65E-03	1,55E-03	1,25E-03	1,54E-03	4,19E-03	8,91E-04	1,74E-03	1,01E-03	2,15E-03	1,18E-03	1,56E-03	1,16E-03
		8	4,33E-03	2,45E-03	1,02E-03	1,30E-03	7,90E-04	2,17E-03	1,39E-03	9,94E-04	2,12E-03	7,47E-02	1,77E-03	2,03E-03	2,71E-03	1,56E-03	7,45E-03	1,55E-03
		9	2,41E-03	4,83E-03	1,29E-03	2,04E-03	1,61E-03	7,98E-04	7,38E-04	1,75E-03	7,16E-03	9,00E-03	5,99E-04	9,26E-04	8,81E-04	2,43E-03	1,39E-02	2,42E-02
		10	3,43E-03	2,12E-03	2,85E-03	2,01E-03	2,29E-03	2,69E-03	7,57E-04	2,04E-03	2,18E-03	8,18E-04	1,60E-03	1,02E-03	2,44E-03	9,55E-04	2,69E-03	3,10E-03
		11	1,42E-03	2,12E-03	1,03E-03	1,88E-03	5,48E-03	4,63E-03	2,71E-03	7,69E-04	1,17E-03	1,83E-03	1,56E-03	2,62E-03	3,65E-03	2,33E-03	2,13E-03	6,90E-04
		12	9,49E-04	9,22E-04	2,66E-03	1,97E-03	1,09E-03	2,74E-03	1,83E-03	8,09E-04	1,15E-03	3,69E-04	1,05E-03	1,29E-03	1,22E-03	2,20E-03	1,37E-03	1,07E-03
		13	3,33E-03	5,99E-03	6,12E-04	3,24E-03	2,08E-03	2,61E-03	5,67E-03	7,75E-04	2,02E-03	3,85E-03	1,78E-03	8,46E-04	2,35E-03	7,60E-04	1,49E-03	5,54E-04
		14	1,98E-03	8,69E-04	1,92E-03	2,14E-03	1,30E-03	2,81E-03	1,09E-03	3,74E-03	6,62E-04	1,08E-03	1,42E-03	2,39E-03	1,57E-03	2,43E-03	6,84E-03	4,76E-03
		15	1,61E-03	8,59E-04	5,31E-03	4,72E-03	1,70E-03	6,72E-04	9,51E-04	1,98E-03	2,75E-03	1,12E-03	6,47E-04	5,08E-04	7,97E-04	5,60E-04	8,49E-04	9,27E-04
		16	1,21E-03	1,15E-03	7,26E-03	1,15E-03	6,68E-03	6,56E-04	1,48E-03	2,13E-03	1,21E-03	1,60E-03	7,70E-04	4,78E-03	1,30E-02	1,10E-03	1,88E-03	7,09E-04
		17	6,96E-04	1,22E-02	3,22E-03	4,52E-03	9,21E-04	4,95E-04	1,40E-03	6,87E-04	3,29E-03	6,09E-04	1,43E-03	1,07E-03	7,06E-04	1,80E-03	7,16E-04	9,40E-03
		18	1,62E-03	6,38E-04	3,53E-03	1,06E-03	1,65E-03	1,74E-03	9,75E-04	1,63E-03	5,72E-04	1,56E-02	5,39E-04	8,92E-04	7,24E-04	1,03E-03	1,34E-03	2,43E-03
		19	1,16E-03	1,11E-03	5,27E-03	2,60E-03	7,10E-04	4,65E-04	4,53E-03	5,80E-04	9,97E-04	6,64E-04	6,87E-04	7,93E-04	1,49E-03	7,97E-04	6,12E-03	4,92E-03
		20	3,38E-03	3,42E-03	7,52E-04	1,33E-03	7,58E-03	5,34E-04	1,96E-03	8,99E-04	4,58E-04	5,01E-04	1,47E-03	5,49E-04	8,53E-04	1,04E-03	5,55E-03	6,70E-03

Πίνακας 4.9 Μέσο τετραγωνικό σφάλμα δεδομένων εκπαίδευσης με την μέθοδο trainGDX.

**Πίνακας 4.10**

Συνολικά αποτελέσματα με τη μέθοδο εκπαίδευσης ανάστροφης διάδοσης των Levenberg-Marquardt (trainLM)

Αρχιτεκτονική	MSE Εκπαίδευσης	MSE Ελέγχου	Εποχές	Χρόνος (sec)	Δεδομένα με σφάλμα >10%
3-5-5-1	2,06E-03	2,25E-04	4	0,188	12
3-5-6-1	8,87E-04	2,37E-04	6	0,203	2
3-5-7-1	7,11E-04	3,73E-05	7	0,218	3
3-5-8-1	9,08E-04	5,85E-05	6	0,203	3
3-5-9-1	6,04E-04	1,82E-04	7	0,218	2
3-5-10-1	8,61E-04	1,43E-04	11	0,281	3
3-5-11-1	6,55E-04	2,03E-04	6	0,219	4
3-5-12-1	4,11E-03	1,86E-04	2	0,172	7
3-5-13-1	8,38E-04	2,45E-04	18	0,375	1
3-5-14-1	9,02E-04	9,34E-05	3	0,187	5
3-5-15-1	1,38E-03	3,03E-04	5	0,218	4
3-5-16-1	1,12E-03	2,61E-04	6	0,234	7
3-5-17-1	8,27E-04	2,47E-04	3	0,188	3
3-5-18-1	8,78E-04	1,90E-04	5	0,219	3
3-5-19-1	5,60E-04	3,88E-04	5	0,218	2
3-5-20-1	1,11E-04	1,72E-04	19	0,437	1
3-6-5-1	7,53E-04	7,77E-05	10	0,265	4
3-6-6-1	6,20E-04	1,73E-04	6	0,219	1
3-6-7-1	9,77E-04	1,79E-04	4	0,188	2
3-6-8-1	6,33E-04	1,70E-04	6	0,218	2
3-6-9-1	4,73E-04	3,11E-04	6	0,219	1
3-6-10-1	3,85E-04	2,27E-04	7	0,219	1
3-6-11-1	4,56E-04	2,12E-04	4	0,203	2
3-6-12-1	9,01E-04	2,83E-04	4	0,203	0
3-6-13-1	2,48E-04	3,05E-04	11	0,296	0
3-6-14-1	7,87E-04	2,83E-04	3	0,187	2
3-6-15-1	5,95E-04	3,83E-04	3	0,187	4
3-6-16-1	1,09E-03	1,68E-04	3	0,203	5
3-6-17-1	6,20E-04	4,52E-04	4	0,202	1
3-6-18-1	7,30E-04	1,90E-04	7	0,250	2
3-6-19-1	7,79E-04	2,26E-04	4	0,219	4
3-6-20-1	4,31E-04	3,16E-04	7	0,281	3
3-7-5-1	6,47E-04	2,35E-04	10	0,265	2
3-7-6-1	4,73E-04	3,01E-04	6	0,219	2
3-7-7-1	6,12E-04	3,16E-04	22	0,437	2
3-7-8-1	6,46E-04	2,53E-04	3	0,171	1
3-7-9-1	7,81E-04	1,46E-04	4	0,187	2
3-7-10-1	5,87E-04	1,25E-04	6	0,219	1
3-7-11-1	6,74E-04	4,36E-04	5	0,203	2
3-7-12-1	7,36E-04	2,79E-04	3	0,187	3
3-7-13-1	3,41E-04	2,53E-04	3	0,187	0
3-7-14-1	4,96E-03	2,72E-04	2	0,172	15

3-7-15-1	5,19E-04	4,22E-04	5	0,234	3
3-7-16-1	3,41E-03	4,99E-04	3	0,203	7
3-7-17-1	1,02E-03	2,25E-04	2	0,172	7
3-7-18-1	4,40E-04	2,15E-04	6	0,265	2
3-7-19-1	8,15E-04	6,29E-05	5	0,249	3
3-7-20-1	3,36E-03	4,09E-04	2	0,188	8
3-8-5-1	1,35E-03	1,51E-04	3	0,171	4
3-8-6-1	6,03E-04	3,39E-04	5	0,187	1
3-8-7-1	1,84E-03	1,35E-04	2	0,156	9
3-8-8-1	1,25E-03	1,38E-04	3	0,187	6
3-8-9-1	9,20E-04	3,91E-04	4	0,187	3
3-8-10-1	5,83E-04	2,29E-04	5	0,202	3
3-8-11-1	4,57E-04	1,51E-04	5	0,234	1
3-8-12-1	6,24E-04	3,20E-04	8	0,265	3
3-8-13-1	4,43E-03	2,65E-04	2	0,187	7
3-8-14-1	1,43E-03	1,51E-04	2	0,188	5
3-8-15-1	4,98E-04	5,33E-04	5	0,234	3
3-8-16-1	5,71E-04	3,44E-04	6	0,265	4
3-8-17-1	1,55E-03	3,70E-04	3	0,203	8
3-8-18-1	1,89E-03	3,47E-04	2	0,203	8
3-8-19-1	6,20E-04	5,06E-04	5	0,249	5
3-8-20-1	5,15E-04	2,69E-04	3	0,219	3
3-9-5-1	1,18E-03	1,48E-04	3	0,171	5
3-9-6-1	8,31E-04	2,75E-04	8	0,234	2
3-9-7-1	6,27E-04	2,11E-04	4	0,172	1
3-9-8-1	1,88E-03	3,07E-04	5	0,219	7
3-9-9-1	4,80E-04	2,06E-04	7	0,250	2
3-9-10-1	3,49E-04	2,48E-04	6	0,234	2
3-9-11-1	6,08E-04	1,78E-04	5	0,234	3
3-9-12-1	4,02E-04	2,44E-04	8	0,281	2
3-9-13-1	2,74E-04	4,79E-04	6	0,203	1
3-9-14-1	3,61E-04	3,06E-04	8	0,297	2
3-9-15-1	1,20E-03	1,29E-04	2	0,203	3
3-9-16-1	6,73E-04	4,55E-04	4	0,234	5
3-9-17-1	8,16E-04	2,69E-04	5	0,265	2
3-9-18-1	1,31E-03	1,77E-04	2	0,218	6
3-9-19-1	1,30E-03	2,74E-04	2	0,218	9
3-9-20-1	4,00E-04	4,03E-04	4	0,281	3
3-10-5-1	6,19E-04	2,59E-04	3	0,172	2
3-10-6-1	6,56E-04	1,31E-04	6	0,234	3
3-10-7-1	1,07E-03	4,57E-04	3	0,187	7
3-10-8-1	9,18E-04	1,30E-04	2	0,172	4
3-10-9-1	3,39E-03	1,61E-04	1	0,156	13
3-10-10-1	2,34E-04	2,03E-04	7	0,250	1
3-10-11-1	6,85E-04	1,30E-04	4	0,218	3
3-10-12-1	4,77E-03	2,92E-04	4	0,218	13
3-10-13-1	1,29E-03	1,14E-04	3	0,202	10
3-10-14-1	6,93E-04	3,09E-04	3	0,218	5

3-10-15-1	1,13E-03	6,30E-04	2	0,203	4
3-10-16-1	4,05E-03	4,46E-04	3	0,218	14
3-10-17-1	8,24E-04	2,76E-04	3	0,218	2
3-10-18-1	2,76E-04	3,23E-04	7	0,327	2
3-10-19-1	1,23E-03	6,57E-04	1	0,218	4
3-10-20-1	6,84E-04	4,42E-04	3	0,266	2
3-11-5-1	2,60E-04	1,20E-04	7	0,219	0
3-11-6-1	3,45E-03	4,09E-04	3	0,172	13
3-11-7-1	1,66E-03	3,33E-04	3	0,187	10
3-11-8-1	8,72E-04	9,98E-05	4	0,203	4
3-11-9-1	1,37E-03	2,03E-04	2	0,140	7
3-11-10-1	3,73E-03	1,38E-04	1	0,156	11
3-11-11-1	8,81E-04	2,14E-04	3	0,202	4
3-11-12-1	7,88E-04	4,12E-04	4	0,218	5
3-11-13-1	7,18E-04	2,93E-04	3	0,218	2
3-11-14-1	1,52E-03	1,70E-04	1	0,187	9
3-11-15-1	1,89E-03	2,75E-04	2	0,203	12
3-11-16-1	3,44E-04	5,12E-04	4	0,250	0
3-11-17-1	1,65E-03	1,09E-04	2	0,234	4
3-11-18-1	2,00E-03	4,08E-04	2	0,250	14
3-11-19-1	2,35E-03	5,72E-04	2	0,234	8
3-11-20-1	5,30E-04	3,34E-04	3	0,281	4
3-12-5-1	1,33E-03	2,27E-04	2	0,171	6
3-12-6-1	7,76E-04	2,99E-04	6	0,234	5
3-12-7-1	6,39E-04	3,72E-04	5	0,234	2
3-12-8-1	4,36E-04	2,49E-04	9	0,296	1
3-12-9-1	5,31E-04	3,96E-04	7	0,250	3
3-12-10-1	1,15E-03	2,20E-04	5	0,249	7
3-12-11-1	5,99E-04	2,74E-04	3	0,202	4
3-12-12-1	3,96E-04	2,91E-04	4	0,234	1
3-12-13-1	6,77E-04	3,02E-04	2	0,203	2
3-12-14-1	1,78E-03	4,10E-04	3	0,218	7
3-12-15-1	3,89E-03	1,35E-04	2	0,234	5
3-12-16-1	1,27E-03	3,16E-04	3	0,281	6
3-12-17-1	2,59E-03	4,81E-04	3	0,234	8
3-12-18-1	5,69E-03	1,89E-04	1	0,234	11
3-12-19-1	1,63E-03	3,91E-04	2	0,266	8
3-12-20-1	8,77E-04	2,32E-04	2	0,249	2
3-13-5-1	5,74E-04	3,28E-04	5	0,219	1
3-13-6-1	4,94E-04	2,82E-04	4	0,203	1
3-13-7-1	1,26E-03	1,90E-04	4	0,203	4
3-13-8-1	3,38E-03	3,98E-04	2	0,171	8
3-13-9-1	5,94E-03	3,01E-04	2	0,203	18
3-13-10-1	3,63E-04	2,25E-04	4	0,234	1
3-13-11-1	2,18E-03	1,20E-04	2	0,203	8
3-13-12-1	1,46E-03	5,16E-04	2	0,203	7
3-13-13-1	1,28E-03	5,62E-04	4	0,265	8
3-13-14-1	1,91E-03	4,35E-04	2	0,202	9

3-13-15-1	2,38E-03	4,32E-04	2	0,249	11
3-13-16-1	1,76E-03	2,94E-04	2	0,250	5
3-13-17-1	3,45E-04	3,14E-04	4	0,312	2
3-13-18-1	1,53E-03	4,75E-04	2	0,280	4
3-13-19-1	6,07E-03	3,83E-04	1	0,234	12
3-13-20-1	2,26E-04	3,70E-04	5	0,343	1
3-14-5-1	5,65E-03	2,02E-04	2	0,171	13
3-14-6-1	2,41E-03	1,86E-04	1	0,171	12
3-14-7-1	2,74E-03	3,57E-04	2	0,187	10
3-14-8-1	2,24E-03	3,04E-04	2	0,187	10
3-14-9-1	1,41E-03	3,90E-04	2	0,203	3
3-14-10-1	7,82E-04	1,61E-04	4	0,234	3
3-14-11-1	1,05E-03	2,71E-04	3	0,234	4
3-14-12-1	1,27E-02	4,06E-04	1	0,203	16
3-14-13-1	6,63E-04	5,60E-04	2	0,234	5
3-14-14-1	2,67E-03	2,84E-04	2	0,234	11
3-14-15-1	6,99E-03	2,52E-04	1	0,234	17
3-11-16-1	1,13E-03	5,74E-04	2	0,265	7
3-14-17-1	1,35E-03	6,54E-04	2	0,296	7
3-14-18-1	1,02E-03	5,66E-04	3	0,328	4
3-14-19-1	3,56E-04	4,02E-04	4	0,359	1
3-14-20-1	1,23E-03	4,07E-04	3	0,374	6
3-15-5-1	3,38E-03	1,47E-04	2	0,171	8
3-15-6-1	5,16E-04	3,63E-04	4	0,218	0
3-15-7-1	4,98E-04	3,15E-04	7	0,266	0
3-15-8-1	1,08E-03	1,92E-04	2	0,187	6
3-15-9-1	5,76E-04	1,91E-04	3	0,219	3
3-15-10-1	1,28E-03	3,09E-04	3	0,234	6
3-15-11-1	4,76E-03	4,01E-04	1	0,202	8
3-15-12-1	7,32E-03	2,39E-04	1	0,218	9
3-15-13-1	3,95E-04	3,33E-04	5	0,328	3
3-15-14-1	7,87E-04	3,35E-04	2	0,234	4
3-15-15-1	7,13E-04	4,67E-04	3	0,312	1
3-15-16-1	3,00E-03	2,84E-04	2	0,281	10
3-15-17-1	4,98E-04	5,97E-04	3	0,328	3
3-15-18-1	3,42E-03	5,25E-04	2	0,296	6
3-15-19-1	1,31E-03	2,90E-04	2	0,343	5
3-15-20-1	3,70E-04	2,74E-04	3	0,390	1
3-16-5-1	3,57E-04	2,53E-04	6	0,249	1
3-16-6-1	2,58E-03	8,88E-05	1	0,156	12
3-16-7-1	2,65E-04	2,23E-04	11	0,343	3
3-16-8-1	4,52E-04	2,90E-04	8	0,312	2
3-16-9-1	3,57E-03	4,16E-04	1	0,171	10
3-16-10-1	8,23E-04	2,14E-04	3	0,234	4
3-16-11-1	5,65E-04	3,62E-04	5	0,297	1
3-16-12-1	4,59E-04	3,18E-04	3	0,265	2
3-16-13-1	4,39E-04	2,17E-04	3	0,281	1
3-16-14-1	2,98E-03	1,94E-04	2	0,265	8



3-16-15-1	4,07E-04	2,60E-04	3	0,312	3
3-16-16-1	7,45E-03	3,22E-04	1	0,249	14
3-16-17-1	1,36E-03	5,94E-04	2	0,312	5
3-16-18-1	5,59E-04	3,11E-04	3	0,374	5
3-16-19-1	1,65E-03	6,36E-04	3	0,343	11
3-16-20-1	2,02E-03	6,25E-04	2	0,343	10
3-17-5-1	1,44E-03	1,52E-04	2	0,171	5
3-17-6-1	1,91E-03	2,62E-04	2	0,188	11
3-17-7-1	6,98E-03	1,32E-04	3	0,219	13
3-17-8-1	3,82E-03	1,83E-04	1	0,218	14
3-17-9-1	1,04E-03	3,64E-04	3	0,530	7
3-17-10-1	6,86E-04	1,85E-04	3	0,608	5
3-17-11-1	2,68E-04	2,56E-04	5	0,327	2
3-17-12-1	2,36E-03	1,32E-04	2	0,265	6
3-17-13-1	1,83E-03	2,63E-04	1	0,265	6
3-17-14-1	1,49E-03	1,48E-04	3	0,687	8
3-17-15-1	7,12E-04	5,62E-04	3	0,328	7
3-17-16-1	2,95E-03	4,22E-04	2	0,671	9
3-17-17-1	4,98E-04	3,31E-04	3	0,359	1
3-17-18-1	8,84E-04	4,40E-04	2	0,359	6
3-17-19-1	1,11E-03	5,02E-04	2	0,390	6
3-17-20-1	1,11E-03	4,83E-04	2	0,405	6
3-18-5-1	1,91E-03	3,09E-04	3	0,219	5
3-18-6-1	1,38E-03	3,70E-04	4	0,561	9
3-18-7-1	6,29E-04	2,88E-04	3	0,219	2
3-18-8-1	1,10E-03	2,62E-04	2	0,218	6
3-18-9-1	2,10E-03	3,67E-04	2	0,202	7
3-18-10-1	1,20E-03	3,57E-04	2	0,593	5
3-18-11-1	1,09E-03	4,01E-04	3	0,297	5
3-18-12-1	2,92E-03	7,35E-04	1	0,250	11
3-18-13-1	6,20E-04	2,61E-04	3	0,265	2
3-18-14-1	1,29E-03	1,75E-04	2	0,328	8
3-18-15-1	3,78E-03	3,33E-04	1	0,296	11
3-18-16-1	7,42E-04	4,64E-04	2	0,266	5
3-18-17-1	5,41E-04	3,42E-04	3	0,390	1
3-18-18-1	1,24E-03	7,47E-04	2	0,358	6
3-18-19-1	1,11E-03	9,23E-05	3	0,437	5
3-18-20-1	5,57E-04	7,93E-04	3	0,484	3
3-19-5-1	1,48E-03	2,64E-04	4	0,593	3
3-19-6-1	1,12E-03	4,12E-04	4	0,609	2
3-19-7-1	2,29E-03	3,15E-04	2	0,546	13
3-19-8-1	2,04E-03	4,20E-04	3	0,609	10
3-19-9-1	5,69E-04	3,75E-04	3	0,624	2
3-19-10-1	2,11E-03	2,38E-04	3	0,265	8
3-19-11-1	8,87E-04	2,49E-04	4	0,250	3
3-19-12-1	5,25E-04	4,34E-04	3	0,327	3
3-19-13-1	4,89E-04	1,12E-04	3	0,265	3
3-19-14-1	1,20E-03	4,58E-04	3	0,390	5

3-19-15-1	1,47E-03	4,56E-04	1	0,312	6
3-19-16-1	1,31E-03	5,13E-04	2	0,359	9
3-19-17-1	2,92E-04	4,56E-04	5	0,499	3
3-19-18-1	4,92E-04	4,42E-04	2	0,374	4
3-19-19-1	1,85E-03	4,27E-04	2	0,437	10
3-19-20-1	5,86E-04	2,41E-04	3	0,514	5
3-20-5-1	4,43E-04	1,66E-04	3	0,203	3
3-20-6-1	9,87E-04	4,25E-04	4	0,608	7
3-20-7-1	5,40E-04	3,20E-04	3	0,624	5
3-20-8-1	2,50E-03	2,67E-04	2	0,234	10
3-20-9-1	1,71E-03	1,23E-04	2	0,265	11
3-20-10-1	3,38E-04	1,73E-04	4	0,686	1
3-20-11-1	7,61E-04	2,27E-04	3	0,702	4
3-20-12-1	2,37E-03	4,33E-04	2	0,687	12
3-20-13-1	6,21E-04	2,86E-04	2	0,312	3
3-20-14-1	4,78E-04	4,29E-04	3	0,733	2
3-20-15-1	4,55E-04	3,03E-04	3	0,406	2
3-20-16-1	4,97E-03	1,07E-04	2	1,155	14
3-20-17-1	8,46E-03	7,87E-04	1	0,296	22
3-20-18-1	2,37E-03	8,71E-04	2	0,437	12
3-20-19-1	2,20E-03	7,27E-04	2	0,453	7
3-20-20-1	1,30E-03	7,01E-04	2	0,468	8

		Αριθμός Νευρώνων 2ου Κρυφού Επιπέδου																
		5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
mse δεδομένων ελέγχου	Αριθμός Νευρώνων 1ου Κρυφού Επιπέδου	5	2,25E-04	2,37E-04	3,73E-05	5,85E-05	1,82E-04	1,43E-04	2,03E-04	1,86E-04	2,45E-04	9,34E-05	3,03E-04	2,61E-04	2,47E-04	1,90E-04	3,88E-04	1,72E-04
		6	7,77E-05	1,73E-04	1,79E-04	1,70E-04	3,11E-04	2,27E-04	2,12E-04	2,83E-04	3,05E-04	2,83E-04	3,83E-04	1,68E-04	4,52E-04	1,90E-04	2,26E-04	3,16E-04
		7	2,35E-04	3,01E-04	3,16E-04	2,53E-04	1,46E-04	1,25E-04	4,36E-04	2,79E-04	2,53E-04	2,72E-04	4,22E-04	4,99E-04	2,25E-04	2,15E-04	6,29E-05	4,09E-04
		8	1,51E-04	3,39E-04	1,35E-04	1,38E-04	3,91E-04	2,29E-04	1,51E-04	3,20E-04	2,65E-04	1,51E-04	5,33E-04	3,44E-04	3,70E-04	3,47E-04	5,06E-04	2,69E-04
		9	1,48E-04	2,75E-04	2,11E-04	3,07E-04	2,06E-04	2,48E-04	1,78E-04	2,44E-04	4,79E-04	3,06E-04	1,29E-04	4,55E-04	2,69E-04	1,77E-04	2,74E-04	4,03E-04
		10	2,59E-04	1,31E-04	4,57E-04	1,30E-04	1,61E-04	2,03E-04	1,30E-04	2,92E-04	1,14E-04	3,09E-04	6,30E-04	4,46E-04	2,76E-04	3,23E-04	6,57E-04	4,42E-04
		11	1,20E-04	4,09E-04	3,33E-04	9,98E-05	2,03E-04	1,38E-04	2,14E-04	4,12E-04	2,93E-04	1,70E-04	2,75E-04	5,12E-04	1,09E-04	4,08E-04	5,72E-04	3,34E-04
		12	2,27E-04	2,99E-04	3,72E-04	2,49E-04	3,96E-04	2,20E-04	2,74E-04	2,91E-04	3,02E-04	4,10E-04	1,35E-04	3,16E-04	4,81E-04	1,89E-04	3,91E-04	2,32E-04
		13	3,28E-04	2,82E-04	1,90E-04	3,98E-04	3,01E-04	2,25E-04	1,20E-04	5,16E-04	5,62E-04	4,35E-04	4,32E-04	2,94E-04	3,14E-04	4,75E-04	3,83E-04	3,70E-04
		14	2,02E-04	1,86E-04	3,57E-04	3,04E-04	3,90E-04	1,61E-04	2,71E-04	4,06E-04	5,60E-04	2,84E-04	2,52E-04	5,74E-04	6,54E-04	5,66E-04	4,02E-04	4,07E-04
		15	1,47E-04	3,63E-04	3,15E-04	1,92E-04	1,91E-04	3,09E-04	4,01E-04	2,39E-04	3,33E-04	3,35E-04	4,67E-04	2,84E-04	5,97E-04	5,25E-04	2,90E-04	2,74E-04
		16	2,53E-04	8,88E-05	2,23E-04	2,90E-04	4,16E-04	2,14E-04	3,62E-04	3,18E-04	2,17E-04	1,94E-04	2,60E-04	3,22E-04	5,94E-04	3,11E-04	6,36E-04	6,25E-04
		17	1,52E-04	2,62E-04	1,32E-04	1,83E-04	3,64E-04	1,85E-04	2,56E-04	1,32E-04	2,63E-04	1,48E-04	5,62E-04	4,22E-04	3,31E-04	4,40E-04	5,02E-04	4,83E-04
		18	3,09E-04	3,70E-04	2,88E-04	2,62E-04	3,67E-04	3,57E-04	4,01E-04	7,35E-04	2,61E-04	1,75E-04	3,33E-04	4,64E-04	3,42E-04	7,47E-04	9,23E-05	7,93E-04
		19	2,64E-04	4,12E-04	3,15E-04	4,20E-04	3,75E-04	2,38E-04	2,49E-04	4,34E-04	1,12E-04	4,58E-04	4,56E-04	5,13E-04	4,56E-04	4,42E-04	4,27E-04	2,41E-04
		20	1,66E-04	4,25E-04	3,20E-04	2,67E-04	1,23E-04	1,73E-04	2,27E-04	4,33E-04	2,86E-04	4,29E-04	3,03E-04	1,07E-04	7,87E-04	8,71E-04	7,27E-04	7,01E-04

Πίνακας 4.11 Μέσο τετραγωνικό σφάλμα δεδομένων ελέγχου με την μέθοδο trainLM.

		Αριθμός Νευρώνων 2ου Κρυφού Επιπέδου																
		5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
Αριθμός Εποχών Εκπαίδευσης	Αριθμός Νευρώνων 1ου Κρυφού Επιπέδου	5	4	6	7	6	7	11	6	2	18	3	5	6	3	5	5	19
		6	10	6	4	6	6	7	4	4	11	3	3	3	4	7	4	7
		7	10	6	22	3	4	6	5	3	3	2	5	3	2	6	5	2
		8	3	5	2	3	4	5	5	8	2	2	5	6	3	2	5	3
		9	3	8	4	5	7	6	5	8	6	8	2	4	5	2	2	4
		10	3	6	3	2	1	7	4	4	3	3	2	3	3	7	1	3
		11	7	3	3	4	2	1	3	4	3	1	2	4	2	2	2	3
		12	2	6	5	9	7	5	3	4	2	3	2	3	3	1	2	2
		13	5	4	4	2	2	4	2	2	4	2	2	2	4	2	1	5
		14	2	1	2	2	2	4	3	1	2	2	1	2	2	3	4	3
		15	2	4	7	2	3	3	1	1	5	2	3	2	3	2	2	3
		16	6	1	11	8	1	3	5	3	3	2	3	1	2	3	3	2
		17	2	2	3	1	3	3	5	2	1	3	3	2	3	2	2	2
		18	3	4	3	2	2	2	3	1	3	2	1	2	3	2	3	3
		19	4	4	2	3	3	3	4	3	3	3	1	2	5	2	2	3
		20	3	4	3	2	2	4	3	2	2	3	3	2	1	2	2	2

Πίνακας 4.12 Αριθμός εποχών που έφθασε κάθε δομή με την μέθοδο trainLM.

Χρόνος Εκπαίδευσης σε sec	Αριθμός Νευρώνων 1ου Κρυφού Επιπέδου	Αριθμός Νευρώνων 2ου Κρυφού Επιπέδου															
		5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
		5	0,188	0,203	0,218	0,203	0,218	0,281	0,219	0,172	0,375	0,187	0,218	0,234	0,188	0,219	0,218
6	0,265	0,219	0,188	0,218	0,219	0,219	0,203	0,203	0,296	0,187	0,187	0,203	0,202	0,25	0,219	0,281	
7	0,265	0,219	0,437	0,171	0,187	0,219	0,203	0,187	0,187	0,172	0,234	0,203	0,172	0,265	0,249	0,188	
8	0,171	0,187	0,156	0,187	0,187	0,202	0,234	0,265	0,187	0,188	0,234	0,265	0,203	0,203	0,249	0,219	
9	0,171	0,234	0,172	0,219	0,25	0,234	0,234	0,281	0,203	0,297	0,203	0,234	0,265	0,218	0,218	0,281	
10	0,172	0,234	0,187	0,172	0,156	0,25	0,218	0,218	0,202	0,218	0,203	0,218	0,218	0,327	0,218	0,266	
11	0,219	0,172	0,187	0,203	0,14	0,156	0,202	0,218	0,218	0,187	0,203	0,25	0,234	0,25	0,234	0,281	
12	0,171	0,234	0,234	0,296	0,25	0,249	0,202	0,234	0,203	0,218	0,234	0,281	0,234	0,234	0,266	0,249	
13	0,219	0,203	0,203	0,171	0,203	0,234	0,203	0,203	0,265	0,202	0,249	0,25	0,312	0,28	0,234	0,343	
14	0,171	0,171	0,187	0,187	0,203	0,234	0,234	0,203	0,234	0,234	0,234	0,265	0,296	0,328	0,359	0,374	
15	0,171	0,218	0,266	0,187	0,219	0,234	0,202	0,218	0,328	0,234	0,312	0,281	0,328	0,296	0,343	0,39	
16	0,249	0,156	0,343	0,312	0,171	0,234	0,297	0,265	0,281	0,265	0,312	0,249	0,312	0,374	0,343	0,343	
17	0,171	0,188	0,219	0,218	0,53	0,608	0,327	0,265	0,265	0,687	0,328	0,671	0,359	0,359	0,39	0,405	
18	0,219	0,561	0,219	0,218	0,202	0,593	0,297	0,25	0,265	0,328	0,296	0,266	0,39	0,358	0,437	0,484	
19	0,593	0,609	0,546	0,609	0,624	0,265	0,25	0,327	0,265	0,39	0,312	0,359	0,499	0,374	0,437	0,514	
20	0,203	0,608	0,624	0,234	0,265	0,686	0,702	0,687	0,312	0,733	0,406	1,155	0,296	0,437	0,453	0,468	

Πίνακας 4.13 Χρόνος εκπαίδευσης κάθε δομής με την μέθοδο trainLM.

		Αριθμός Νευρώνων 2ου Κρυφού Επιπέδου																
		5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
mse δεδομένων εκπαίδευσης	Αριθμός Νευρώνων 1ου Κρυφού Επιπέδου	5	2,06E-03	8,87E-04	7,11E-04	9,08E-04	6,04E-04	8,61E-04	6,55E-04	4,11E-03	8,38E-04	9,02E-04	1,38E-03	1,12E-03	8,27E-04	8,78E-04	5,60E-04	1,11E-04
		6	7,53E-04	6,20E-04	9,77E-04	6,33E-04	4,73E-04	3,85E-04	4,56E-04	9,01E-04	2,48E-04	7,87E-04	5,95E-04	1,09E-03	6,20E-04	7,30E-04	7,79E-04	4,31E-04
		7	6,47E-04	4,73E-04	6,12E-04	6,46E-04	7,81E-04	5,87E-04	6,74E-04	7,36E-04	3,41E-04	4,96E-03	5,19E-04	3,41E-03	1,02E-03	4,40E-04	8,15E-04	3,36E-03
		8	1,35E-03	6,03E-04	1,84E-03	1,25E-03	9,20E-04	5,83E-04	4,57E-04	6,24E-04	4,43E-03	1,43E-03	4,98E-04	5,71E-04	1,55E-03	1,89E-03	6,20E-04	5,15E-04
		9	1,18E-03	8,31E-04	6,27E-04	1,88E-03	4,80E-04	3,49E-04	6,08E-04	4,02E-04	2,74E-04	3,61E-04	1,20E-03	6,73E-04	8,16E-04	1,31E-03	1,30E-03	4,00E-04
		10	6,19E-04	6,56E-04	1,07E-03	9,18E-04	3,39E-03	2,34E-04	6,85E-04	4,77E-03	1,29E-03	6,93E-04	1,13E-03	4,05E-03	8,24E-04	2,76E-04	1,23E-03	6,84E-04
		11	2,60E-04	3,45E-03	1,66E-03	8,72E-04	1,37E-03	3,73E-03	8,81E-04	7,88E-04	7,18E-04	1,52E-03	1,89E-03	3,44E-04	1,65E-03	2,00E-03	2,35E-03	5,30E-04
		12	1,33E-03	7,76E-04	6,39E-04	4,36E-04	5,31E-04	1,15E-03	5,99E-04	3,96E-04	6,77E-04	1,78E-03	3,89E-03	1,27E-03	2,59E-03	5,69E-03	1,63E-03	8,77E-04
		13	5,74E-04	4,94E-04	1,26E-03	3,38E-03	5,94E-03	3,63E-04	2,18E-03	1,46E-03	1,28E-03	1,91E-03	2,38E-03	1,76E-03	3,45E-04	1,53E-03	6,07E-03	2,26E-04
		14	5,65E-03	2,41E-03	2,74E-03	2,24E-03	1,41E-03	7,82E-04	1,05E-03	1,27E-02	6,63E-04	2,67E-03	6,99E-03	1,13E-03	1,35E-03	1,02E-03	3,56E-04	1,23E-03
		15	3,38E-03	5,16E-04	4,98E-04	1,08E-03	5,76E-04	1,28E-03	4,76E-03	7,32E-03	3,95E-04	7,87E-04	7,13E-04	3,00E-03	4,98E-04	3,42E-03	1,31E-03	3,70E-04
		16	3,57E-04	2,58E-03	2,65E-04	4,52E-04	3,57E-03	8,23E-04	5,65E-04	4,59E-04	4,39E-04	2,98E-03	4,07E-04	7,45E-03	1,36E-03	5,59E-04	1,65E-03	2,02E-03
		17	1,44E-03	1,91E-03	6,98E-03	3,82E-03	1,04E-03	6,86E-04	2,68E-04	2,36E-03	1,83E-03	1,49E-03	7,12E-04	2,95E-03	4,98E-04	8,84E-04	1,11E-03	1,11E-03
		18	1,91E-03	1,38E-03	6,29E-04	1,10E-03	2,10E-03	1,20E-03	1,09E-03	2,92E-03	6,20E-04	1,29E-03	3,78E-03	7,42E-04	5,41E-04	1,24E-03	1,11E-03	5,57E-04
		19	1,48E-03	1,12E-03	2,29E-03	2,04E-03	5,69E-04	2,11E-03	8,87E-04	5,25E-04	4,89E-04	1,20E-03	1,47E-03	1,31E-03	2,92E-04	4,92E-04	1,85E-03	5,86E-04
		20	4,43E-04	9,87E-04	5,40E-04	2,50E-03	1,71E-03	3,38E-04	7,61E-04	2,37E-03	6,21E-04	4,78E-04	4,55E-04	4,97E-03	8,46E-03	2,37E-03	2,20E-03	1,30E-03

Πίνακας 4.14 Μέσο τετραγωνικό σφάλμα δεδομένων εκπαίδευσης με την μέθοδο trainLM.

**Πίνακας 4.15**

Συνολικά αποτελέσματα με τη μέθοδο εκπαίδευσης συζυγών κλίσεων "scaled conjugate gradient" (trainscg)

Αρχιτεκτονική	MSE Εκπαίδευσης	MSE Ελέγχου	Εποχές	Χρόνος (sec)	Δεδομένα με σφάλμα >10%
3-5-5-1	1,18E-03	1,57E-04	31	0,047	5
3-5-6-1	5,86E-03	2,38E-04	7	0,015	16
3-5-7-1	3,28E-03	3,96E-04	13	0,031	10
3-5-8-1	3,95E-03	2,03E-04	7	0,031	16
3-5-9-1	2,72E-03	4,06E-04	10	0,015	13
3-5-10-1	1,76E-03	3,68E-04	24	0,047	9
3-5-11-1	9,40E-04	3,11E-04	40	0,047	3
3-5-12-1	1,40E-03	3,34E-04	17	0,016	6
3-5-13-1	1,29E-03	2,16E-04	19	0,031	8
3-5-14-1	2,09E-03	2,45E-04	18	0,031	10
3-5-15-1	1,27E-03	3,23E-04	27	0,057	5
3-5-16-1	2,47E-03	2,39E-04	9	0,021	10
3-5-17-1	7,21E-03	1,50E-04	5	0,016	18
3-5-18-1	2,30E-03	2,92E-04	14	0,016	11
3-5-19-1	6,16E-03	2,38E-04	6	0,015	12
3-5-20-1	1,74E-03	2,27E-04	25	0,047	10
3-6-5-1	1,31E-03	1,66E-04	32	0,047	4
3-6-6-1	1,14E-03	3,88E-04	23	0,031	7
3-6-7-1	2,67E-03	1,59E-04	19	0,031	11
3-6-8-1	2,63E-03	3,35E-04	14	0,031	9
3-6-9-1	1,43E-03	1,02E-04	16	0,047	6
3-6-10-1	4,08E-03	2,38E-04	13	0,031	13
3-6-11-1	1,35E-03	9,87E-05	15	0,031	4
3-6-12-1	2,13E-03	1,44E-04	20	0,036	7
3-6-13-1	1,84E-03	2,91E-04	18	0,031	9
3-6-14-1	4,07E-03	2,57E-04	4	0,016	14
3-6-15-1	1,05E-03	2,37E-04	21	0,032	5
3-6-16-1	8,22E-04	4,08E-04	38	0,047	6
3-6-17-1	3,32E-03	4,69E-04	8	0,015	12
3-6-18-1	2,45E-03	4,22E-04	12	0,016	10
3-6-19-1	6,59E-03	2,80E-04	6	0,016	14
3-6-20-1	2,44E-03	4,27E-04	12	0,031	11
3-7-5-1	3,95E-03	1,44E-04	5	0,015	12
3-7-6-1	1,52E-03	2,50E-04	29	0,046	7
3-7-7-1	3,42E-03	2,94E-04	8	0,016	11
3-7-8-1	5,98E-03	2,72E-04	6	0,016	19
3-7-9-1	6,81E-03	2,18E-04	5	0,016	17
3-7-10-1	1,90E-03	4,62E-04	11	0,016	7
3-7-11-1	2,41E-03	3,62E-04	10	0,015	14
3-7-12-1	1,21E-03	2,40E-04	30	0,047	4
3-7-13-1	1,13E-03	2,33E-04	18	0,032	8
3-7-14-1	2,30E-03	3,07E-04	16	0,016	12

3-7-15-1	1,04E-03	1,62E-04	30	0,063	3
3-7-16-1	2,93E-03	1,48E-04	13	0,031	10
3-7-17-1	1,38E-03	3,25E-04	30	0,047	10
3-7-18-1	1,34E-03	1,61E-04	22	0,032	7
3-7-19-1	1,82E-03	2,15E-04	10	0,016	6
3-7-20-1	1,08E-03	1,94E-04	28	0,047	4
3-8-5-1	5,68E-04	1,45E-04	63	0,080	1
3-8-6-1	1,76E-03	2,02E-04	11	0,016	5
3-8-7-1	4,55E-03	8,92E-05	3	0,015	16
3-8-8-1	2,61E-03	3,22E-04	12	0,016	11
3-8-9-1	5,29E-04	1,69E-04	40	0,062	1
3-8-10-1	2,35E-03	2,85E-04	17	0,016	12
3-8-11-1	8,93E-04	2,83E-04	36	0,047	2
3-8-12-1	3,35E-03	2,62E-04	8	0,016	10
3-8-13-1	2,11E-03	2,00E-04	15	0,016	5
3-8-14-1	4,01E-03	8,36E-05	4	0,000	18
3-8-15-1	2,14E-03	4,54E-04	11	0,016	8
3-8-16-1	2,61E-03	3,60E-04	15	0,031	12
3-8-17-1	8,88E-04	2,88E-04	26	0,049	4
3-8-18-1	7,21E-04	1,70E-04	19	0,047	2
3-8-19-1	4,95E-03	4,64E-04	7	0,015	12
3-8-20-1	8,91E-04	3,19E-04	32	0,047	3
3-9-5-1	5,05E-03	3,10E-04	4	0,016	13
3-9-6-1	1,14E-03	4,54E-05	18	0,031	1
3-9-7-1	1,76E-03	1,13E-04	10	0,016	6
3-9-8-1	1,59E-03	2,05E-04	20	0,031	8
3-9-9-1	7,05E-04	3,03E-04	52	0,094	4
3-9-10-1	1,48E-03	2,59E-04	20	0,031	10
3-9-11-1	1,37E-03	2,06E-04	10	0,037	5
3-9-12-1	1,39E-03	2,61E-04	9	0,016	5
3-9-13-1	1,68E-03	9,52E-05	18	0,031	8
3-9-14-1	1,15E-03	1,42E-04	13	0,031	4
3-9-15-1	1,13E-03	1,80E-04	13	0,016	3
3-9-16-1	1,45E-03	3,58E-04	22	0,031	11
3-9-17-1	4,00E-03	2,37E-04	10	0,031	13
3-9-18-1	6,59E-04	3,76E-04	29	0,062	3
3-9-19-1	1,19E-03	5,19E-04	16	0,032	7
3-9-20-1	5,38E-03	5,04E-04	8	0,015	16
3-10-5-1	2,92E-03	2,79E-04	11	0,031	14
3-10-6-1	1,54E-03	2,45E-04	24	0,047	6
3-10-7-1	2,68E-03	2,37E-04	13	0,031	12
3-10-8-1	1,60E-03	2,92E-04	12	0,032	8
3-10-9-1	2,37E-03	1,25E-04	13	0,031	11
3-10-10-1	8,42E-04	2,04E-04	16	0,027	5
3-10-11-1	3,83E-03	3,53E-04	11	0,032	14
3-10-12-1	1,11E-03	3,31E-04	21	0,031	3
3-10-13-1	1,93E-03	2,00E-04	14	0,015	10
3-10-14-1	2,67E-03	2,72E-04	7	0,016	13



3-10-15-1	5,69E-04	2,66E-04	44	0,047	2
3-10-16-1	1,22E-03	1,30E-04	10	0,016	4
3-10-17-1	1,48E-03	4,76E-04	12	0,032	8
3-10-18-1	1,22E-03	3,69E-04	15	0,031	6
3-10-19-1	2,52E-03	2,23E-04	14	0,016	11
3-10-20-1	1,23E-03	2,98E-04	23	0,031	4
3-11-5-1	8,26E-04	3,68E-04	23	0,031	3
3-11-6-1	2,32E-03	1,97E-04	12	0,031	11
3-11-7-1	8,04E-04	2,53E-04	41	0,078	1
3-11-8-1	1,32E-03	2,92E-04	12	0,031	7
3-11-9-1	1,70E-03	3,33E-04	24	0,047	11
3-11-10-1	6,88E-04	3,26E-04	31	0,047	5
3-11-11-1	1,17E-03	2,12E-04	18	0,031	9
3-11-12-1	1,06E-03	2,77E-04	13	0,031	5
3-11-13-1	1,55E-03	7,89E-05	10	0,016	9
3-11-14-1	1,13E-03	1,36E-04	24	0,062	5
3-11-15-1	8,01E-03	5,36E-04	5	0,015	18
3-11-16-1	5,67E-04	4,11E-04	32	0,047	2
3-11-17-1	5,09E-03	2,38E-04	5	0,015	16
3-11-18-1	2,89E-03	1,45E-04	8	0,015	10
3-11-19-1	8,40E-04	2,95E-04	39	0,062	5
3-11-20-1	1,27E-03	2,01E-04	12	0,031	9
3-12-5-1	2,42E-03	2,90E-04	7	0,015	10
3-12-6-1	3,20E-03	2,74E-04	17	0,031	7
3-12-7-1	1,93E-03	2,97E-04	21	0,031	12
3-12-8-1	1,35E-03	1,65E-04	15	0,031	4
3-12-9-1	2,29E-03	1,20E-04	10	0,031	9
3-12-10-1	1,24E-03	3,89E-04	12	0,031	7
3-12-11-1	1,74E-03	1,94E-04	12	0,016	7
3-12-12-1	1,22E-03	2,94E-04	10	0,031	5
3-12-13-1	1,14E-03	3,85E-04	13	0,015	7
3-12-14-1	1,01E-03	3,09E-04	14	0,031	4
3-12-15-1	1,78E-03	2,87E-04	8	0,016	11
3-12-16-1	4,28E-03	4,06E-04	8	0,016	15
3-12-17-1	9,56E-04	3,53E-04	26	0,032	2
3-12-18-1	5,81E-04	2,96E-04	55	0,079	0
3-12-19-1	1,27E-03	3,88E-04	19	0,031	7
3-12-20-1	1,34E-03	4,95E-04	12	0,047	8
3-13-5-1	1,54E-03	1,50E-04	24	0,032	8
3-13-6-1	2,92E-03	2,34E-04	8	0,032	12
3-13-7-1	1,30E-03	6,33E-05	13	0,031	4
3-13-8-1	3,60E-03	2,92E-04	8	0,032	14
3-13-9-1	8,32E-04	3,19E-04	28	0,047	1
3-13-10-1	1,40E-03	2,10E-04	10	0,016	6
3-13-11-1	6,89E-04	1,98E-04	45	0,062	3
3-13-12-1	1,07E-03	3,37E-04	23	0,046	4
3-13-13-1	7,52E-04	4,21E-04	15	0,031	4
3-13-14-1	1,17E-03	1,47E-04	25	0,032	6

3-13-15-1	1,84E-03	2,63E-04	16	0,016	9
3-13-16-1	6,47E-04	2,96E-04	44	0,062	2
3-13-17-1	7,20E-04	3,14E-04	35	0,062	5
3-13-18-1	7,76E-04	4,93E-04	23	0,047	3
3-13-19-1	4,25E-04	4,47E-04	48	0,078	2
3-13-20-1	1,79E-03	5,41E-04	11	0,017	8
3-14-5-1	5,45E-04	1,34E-04	30	0,047	1
3-14-6-1	2,83E-03	2,36E-04	7	0,015	14
3-14-7-1	1,26E-03	2,09E-04	18	0,015	6
3-14-8-1	1,39E-03	1,84E-04	12	0,015	9
3-14-9-1	1,10E-03	1,65E-04	24	0,047	5
3-14-10-1	7,87E-04	3,11E-04	31	0,047	3
3-14-11-1	7,64E-04	4,49E-04	19	0,031	1
3-14-12-1	2,12E-03	2,26E-04	12	0,015	12
3-14-13-1	6,25E-04	2,55E-04	46	0,062	2
3-14-14-1	1,14E-03	1,77E-04	10	0,031	4
3-14-15-1	9,69E-04	2,68E-04	36	0,047	4
3-11-16-1	5,23E-04	3,05E-04	28	0,047	2
3-14-17-1	1,21E-03	3,89E-04	13	0,031	4
3-14-18-1	1,22E-03	1,74E-04	10	0,031	6
3-14-19-1	3,20E-03	4,95E-04	7	0,015	15
3-14-20-1	3,94E-04	3,46E-04	37	0,062	1
3-15-5-1	9,39E-04	2,28E-04	15	0,031	3
3-15-6-1	8,29E-03	9,02E-05	2	0,016	19
3-15-7-1	3,87E-04	8,32E-05	49	0,062	2
3-15-8-1	1,64E-03	3,64E-04	10	0,016	5
3-15-9-1	2,55E-03	2,13E-04	10	0,015	9
3-15-10-1	2,39E-03	3,54E-04	7	0,015	9
3-15-11-1	1,14E-03	2,26E-04	11	0,016	6
3-15-12-1	6,53E-04	2,80E-04	17	0,031	1
3-15-13-1	1,18E-03	3,30E-04	14	0,015	6
3-15-14-1	6,86E-04	2,51E-04	35	0,047	2
3-15-15-1	7,00E-04	3,32E-04	23	0,031	3
3-15-16-1	1,35E-03	1,29E-04	16	0,031	3
3-15-17-1	1,87E-03	2,77E-04	9	0,016	7
3-15-18-1	1,67E-03	3,77E-04	13	0,031	8
3-15-19-1	9,61E-04	2,95E-04	17	0,032	4
3-15-20-1	1,04E-03	4,46E-04	22	0,047	5
3-16-5-1	8,35E-04	2,24E-04	27	0,047	3
3-16-6-1	1,25E-03	1,53E-04	17	0,032	7
3-16-7-1	1,45E-03	2,13E-04	22	0,031	10
3-16-8-1	1,22E-03	1,24E-04	16	0,047	5
3-16-9-1	7,62E-04	2,29E-04	22	0,046	3
3-16-10-1	1,31E-03	3,20E-04	16	0,016	6
3-16-11-1	1,52E-03	4,40E-04	13	0,031	5
3-16-12-1	8,09E-03	2,13E-04	5	0,015	18
3-16-13-1	7,69E-04	1,56E-04	26	0,031	3
3-16-14-1	7,63E-04	4,04E-04	26	0,046	2

3-16-15-1	3,22E-03	1,75E-04	9	0,016	11
3-16-16-1	1,00E-03	2,36E-04	15	0,031	4
3-16-17-1	7,28E-04	2,71E-04	26	0,046	3
3-16-18-1	5,84E-04	3,64E-04	38	0,062	0
3-16-19-1	2,09E-03	3,02E-04	8	0,016	8
3-16-20-1	1,39E-03	6,71E-04	12	0,172	7
3-17-5-1	2,18E-03	2,18E-04	9	0,015	13
3-17-6-1	3,97E-03	2,77E-04	7	0,015	13
3-17-7-1	1,18E-03	2,22E-04	30	0,047	5
3-17-8-1	2,04E-03	2,36E-04	14	0,031	11
3-17-9-1	1,64E-03	1,84E-04	14	0,031	9
3-17-10-1	2,89E-03	1,74E-04	7	0,016	11
3-17-11-1	9,10E-04	2,72E-04	28	0,063	5
3-17-12-1	8,77E-04	2,89E-04	24	0,041	6
3-17-13-1	6,66E-04	1,93E-04	17	0,033	1
3-17-14-1	2,51E-03	4,25E-04	11	0,032	11
3-17-15-1	1,15E-03	2,68E-04	13	0,032	6
3-17-16-1	7,33E-04	2,58E-04	15	0,031	2
3-17-17-1	1,66E-03	5,95E-04	13	0,016	7
3-17-18-1	2,07E-03	6,45E-04	10	0,032	9
3-17-19-1	1,13E-03	2,15E-04	23	0,047	5
3-17-20-1	2,20E-03	3,33E-04	12	0,031	12
3-18-5-1	3,06E-03	2,02E-04	6	0,000	14
3-18-6-1	5,35E-04	1,97E-04	50	0,063	2
3-18-7-1	5,51E-04	1,93E-04	31	0,047	1
3-18-8-1	5,51E-04	1,01E-04	29	0,046	2
3-18-9-1	3,98E-04	2,83E-04	50	0,062	0
3-18-10-1	2,57E-03	3,06E-04	7	0,033	12
3-18-11-1	6,83E-04	3,39E-04	30	0,047	2
3-18-12-1	1,61E-03	2,69E-04	14	0,031	10
3-18-13-1	1,28E-03	2,35E-04	13	0,016	8
3-18-14-1	6,55E-04	4,35E-04	30	0,047	4
3-18-15-1	7,68E-04	6,40E-05	19	0,031	3
3-18-16-1	7,53E-04	5,07E-04	18	0,032	4
3-18-17-1	5,63E-04	4,35E-04	28	0,047	2
3-18-18-1	6,54E-04	6,87E-04	33	0,062	3
3-18-19-1	3,02E-03	2,56E-04	9	0,031	11
3-18-20-1	2,37E-03	4,67E-04	15	0,031	9
3-19-5-1	1,39E-03	3,16E-04	16	0,031	6
3-19-6-1	2,14E-03	1,21E-04	6	0,016	9
3-19-7-1	3,75E-03	2,74E-04	7	0,031	14
3-19-8-1	1,21E-03	2,25E-04	12	0,031	5
3-19-9-1	7,99E-03	2,90E-04	5	0,015	14
3-19-10-1	6,18E-04	1,88E-04	54	0,129	1
3-19-11-1	1,12E-03	1,68E-04	19	0,031	7
3-19-12-1	5,28E-04	5,69E-04	30	0,047	2
3-19-13-1	1,22E-03	4,06E-04	12	0,016	5
3-19-14-1	3,33E-03	4,73E-04	8	0,015	11

3-19-15-1	1,06E-03	2,39E-04	20	0,047	1
3-19-16-1	5,75E-04	4,77E-04	36	0,062	3
3-19-17-1	8,26E-04	5,61E-04	33	0,047	6
3-19-18-1	7,27E-04	4,08E-04	18	0,031	1
3-19-19-1	4,92E-04	5,05E-04	38	0,078	0
3-19-20-1	4,51E-04	4,28E-04	70	0,109	2
3-20-5-1	6,55E-04	1,86E-04	23	0,041	2
3-20-6-1	1,42E-03	1,70E-04	18	0,034	5
3-20-7-1	1,10E-03	2,19E-04	19	0,031	2
3-20-8-1	2,58E-03	2,24E-04	9	0,015	14
3-20-9-1	1,08E-03	2,72E-04	15	0,032	5
3-20-10-1	2,79E-03	2,28E-04	10	0,031	10
3-20-11-1	2,36E-03	2,88E-04	8	0,031	10
3-20-12-1	2,53E-03	4,50E-04	8	0,016	9
3-20-13-1	8,59E-04	4,75E-04	17	0,031	2
3-20-14-1	6,59E-04	2,90E-04	24	0,047	1
3-20-15-1	5,67E-04	3,30E-04	64	0,078	1
3-20-16-1	1,39E-03	2,96E-04	10	0,025	6
3-20-17-1	7,87E-03	2,33E-04	4	0,016	16
3-20-18-1	3,52E-04	4,55E-04	56	0,078	0
3-20-19-1	2,83E-03	4,99E-04	9	0,031	13
3-20-20-1	9,92E-04	6,71E-04	14	0,016	5

		Αριθμός Νευρώνων 2ου Κρυφού Επιπέδου																
		5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
mse δεδομένων ελέγχου	Αριθμός Νευρώνων 1ου Κρυφού Επιπέδου	5	1,57E-04	2,38E-04	3,96E-04	2,03E-04	4,06E-04	3,68E-04	3,11E-04	3,34E-04	2,16E-04	2,45E-04	3,23E-04	2,39E-04	1,50E-04	2,92E-04	2,38E-04	2,27E-04
		6	1,66E-04	3,88E-04	1,59E-04	3,35E-04	1,02E-04	2,38E-04	9,87E-05	1,44E-04	2,91E-04	2,57E-04	2,37E-04	4,08E-04	4,69E-04	4,22E-04	2,80E-04	4,27E-04
		7	1,44E-04	2,50E-04	2,94E-04	2,72E-04	2,18E-04	4,62E-04	3,62E-04	2,40E-04	2,33E-04	3,07E-04	1,62E-04	1,48E-04	3,25E-04	1,61E-04	2,15E-04	1,94E-04
		8	1,45E-04	2,02E-04	8,92E-05	3,22E-04	1,69E-04	2,85E-04	2,83E-04	2,62E-04	2,00E-04	8,36E-05	4,54E-04	3,60E-04	2,88E-04	1,70E-04	4,64E-04	3,19E-04
		9	3,10E-04	4,54E-05	1,13E-04	2,05E-04	3,03E-04	2,59E-04	2,06E-04	2,61E-04	9,52E-05	1,42E-04	1,80E-04	3,58E-04	2,37E-04	3,76E-04	5,19E-04	5,04E-04
		10	2,79E-04	2,45E-04	2,37E-04	2,92E-04	1,25E-04	2,04E-04	3,53E-04	3,31E-04	2,00E-04	2,72E-04	2,66E-04	1,30E-04	4,76E-04	3,69E-04	2,23E-04	2,98E-04
		11	3,68E-04	1,97E-04	2,53E-04	2,92E-04	3,33E-04	3,26E-04	2,12E-04	2,77E-04	7,89E-05	1,36E-04	5,36E-04	4,11E-04	2,38E-04	1,45E-04	2,95E-04	2,01E-04
		12	2,90E-04	2,74E-04	2,97E-04	1,65E-04	1,20E-04	3,89E-04	1,94E-04	2,94E-04	3,85E-04	3,09E-04	2,87E-04	4,06E-04	3,53E-04	2,96E-04	3,88E-04	4,95E-04
		13	1,50E-04	2,34E-04	6,33E-05	2,92E-04	3,19E-04	2,10E-04	1,98E-04	3,37E-04	4,21E-04	1,47E-04	2,63E-04	2,96E-04	3,14E-04	4,93E-04	4,47E-04	5,41E-04
		14	1,34E-04	2,36E-04	2,09E-04	1,84E-04	1,65E-04	3,11E-04	4,49E-04	2,26E-04	2,55E-04	1,77E-04	2,68E-04	3,05E-04	3,89E-04	1,74E-04	4,95E-04	3,46E-04
		15	2,28E-04	9,02E-05	8,32E-05	3,64E-04	2,13E-04	3,54E-04	2,26E-04	2,80E-04	3,30E-04	2,51E-04	3,32E-04	1,29E-04	2,77E-04	3,77E-04	2,95E-04	4,46E-04
		16	2,24E-04	1,53E-04	2,13E-04	1,24E-04	2,29E-04	3,20E-04	4,40E-04	2,13E-04	1,56E-04	4,04E-04	1,75E-04	2,36E-04	2,71E-04	3,64E-04	3,02E-04	6,71E-04
		17	2,18E-04	2,77E-04	2,22E-04	2,36E-04	1,84E-04	1,74E-04	2,72E-04	2,89E-04	1,93E-04	4,25E-04	2,68E-04	2,58E-04	5,95E-04	6,45E-04	2,15E-04	3,33E-04
		18	2,02E-04	1,97E-04	1,93E-04	1,01E-04	2,83E-04	3,06E-04	3,39E-04	2,69E-04	2,35E-04	4,35E-04	6,40E-05	5,07E-04	4,35E-04	6,87E-04	2,56E-04	4,67E-04
		19	3,16E-04	1,21E-04	2,74E-04	2,25E-04	2,90E-04	1,88E-04	1,68E-04	5,69E-04	4,06E-04	4,73E-04	2,39E-04	4,77E-04	5,61E-04	4,08E-04	5,05E-04	4,28E-04
		20	1,86E-04	1,70E-04	2,19E-04	2,24E-04	2,72E-04	2,28E-04	2,88E-04	4,50E-04	4,75E-04	2,90E-04	3,30E-04	2,96E-04	2,33E-04	4,55E-04	4,99E-04	6,71E-04

Πίνακας 4.16 Μέσο τετραγωνικό σφάλμα δεδομένων ελέγχου με την μέθοδο trainSCG.

		Αριθμός Νευρώνων 2ου Κρυφού Επιπέδου																
		5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
Αριθμός Εποχών Εκπαίδευσης	Αριθμός Νευρώνων 1ου Κρυφού Επιπέδου	5	31	7	13	7	10	24	40	17	19	18	27	9	5	14	6	25
		6	32	23	19	14	16	13	15	20	18	4	21	38	8	12	6	12
		7	5	29	8	6	5	11	10	30	18	16	30	13	30	22	10	28
		8	63	11	3	12	40	17	36	8	15	4	11	15	26	19	7	32
		9	4	18	10	20	52	20	10	9	18	13	13	22	10	29	16	8
		10	11	24	13	12	13	16	11	21	14	7	44	10	12	15	14	23
		11	23	12	41	12	24	31	18	13	10	24	5	32	5	8	39	12
		12	7	17	21	15	10	12	12	10	13	14	8	8	26	55	19	12
		13	24	8	13	8	28	10	45	23	15	25	16	44	35	23	48	11
		14	30	7	18	12	24	31	19	12	46	10	36	28	13	10	7	37
		15	15	2	49	10	10	7	11	17	14	35	23	16	9	13	17	22
		16	27	17	22	16	22	16	13	5	26	26	9	15	26	38	8	12
		17	9	7	30	14	14	7	28	24	17	11	13	15	13	10	23	12
		18	6	50	31	29	50	7	30	14	13	30	19	18	28	33	9	15
		19	16	6	7	12	5	54	19	30	12	8	20	36	33	18	38	70
		20	23	18	19	9	15	10	8	8	17	24	64	10	4	56	9	14

Πίνακας 4.17 Αριθμός εποχών που έφθασε κάθε δομή με την μέθοδο trainSCG.

Χρόνος Εκπαίδευσης σε sec	Αριθμός Νευρώνων 1ου Κρυφού Επιπέδου	Αριθμός Νευρώνων 2ου Κρυφού Επιπέδου															
		5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
		5	0,047	0,015	0,031	0,031	0,015	0,047	0,047	0,016	0,031	0,031	0,057	0,021	0,016	0,016	0,015
6	0,047	0,031	0,031	0,031	0,047	0,031	0,031	0,036	0,031	0,016	0,032	0,047	0,015	0,016	0,016	0,031	
7	0,015	0,046	0,016	0,016	0,016	0,016	0,015	0,047	0,032	0,016	0,063	0,031	0,047	0,032	0,016	0,047	
8	0,08	0,016	0,015	0,016	0,062	0,016	0,047	0,016	0,016	0	0,016	0,031	0,049	0,047	0,015	0,047	
9	0,016	0,031	0,016	0,031	0,094	0,031	0,037	0,016	0,031	0,031	0,016	0,031	0,031	0,062	0,032	0,015	
10	0,031	0,047	0,031	0,032	0,031	0,027	0,032	0,031	0,015	0,016	0,047	0,016	0,032	0,031	0,016	0,031	
11	0,031	0,031	0,078	0,031	0,047	0,047	0,031	0,031	0,016	0,062	0,015	0,047	0,015	0,015	0,062	0,031	
12	0,015	0,031	0,031	0,031	0,031	0,031	0,031	0,016	0,031	0,015	0,031	0,016	0,016	0,032	0,079	0,031	
13	0,032	0,032	0,031	0,032	0,047	0,016	0,062	0,046	0,031	0,032	0,016	0,062	0,062	0,047	0,078	0,017	
14	0,047	0,015	0,015	0,015	0,047	0,047	0,031	0,015	0,062	0,031	0,047	0,047	0,031	0,031	0,015	0,062	
15	0,031	0,016	0,062	0,016	0,015	0,015	0,016	0,031	0,015	0,047	0,031	0,031	0,016	0,031	0,032	0,047	
16	0,047	0,032	0,031	0,047	0,046	0,016	0,031	0,015	0,031	0,046	0,016	0,031	0,046	0,062	0,016	0,172	
17	0,015	0,015	0,047	0,031	0,031	0,016	0,063	0,041	0,033	0,032	0,032	0,031	0,016	0,032	0,047	0,031	
18	0	0,063	0,047	0,046	0,062	0,033	0,047	0,031	0,016	0,047	0,031	0,032	0,047	0,062	0,031	0,031	
19	0,031	0,016	0,031	0,031	0,015	0,129	0,031	0,047	0,016	0,015	0,047	0,062	0,047	0,031	0,078	0,109	
20	0,041	0,034	0,031	0,015	0,032	0,031	0,031	0,016	0,031	0,047	0,078	0,025	0,016	0,078	0,031	0,016	

Πίνακας 4.18 Χρόνος εκπαίδευσης κάθε δομής με την μέθοδο trainSCG.

		Αριθμός Νευρώνων 2ου Κρυφού Επιπέδου																
		5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
mse δεδομένων εκπαίδευσης	Αριθμός Νευρώνων 1ου Κρυφού Επιπέδου	5	1,18E-03	5,86E-03	3,28E-03	3,95E-03	2,72E-03	1,76E-03	9,40E-04	1,40E-03	1,29E-03	2,09E-03	1,27E-03	2,47E-03	7,21E-03	2,30E-03	6,16E-03	1,74E-03
		6	1,31E-03	1,14E-03	2,67E-03	2,63E-03	1,43E-03	4,08E-03	1,35E-03	2,13E-03	1,84E-03	4,07E-03	1,05E-03	8,22E-04	3,32E-03	2,45E-03	6,59E-03	2,44E-03
		7	3,95E-03	1,52E-03	3,42E-03	5,98E-03	6,81E-03	1,90E-03	2,41E-03	1,21E-03	1,13E-03	2,30E-03	1,04E-03	2,93E-03	1,38E-03	1,34E-03	1,82E-03	1,08E-03
		8	5,68E-04	1,76E-03	4,55E-03	2,61E-03	5,29E-04	2,35E-03	8,93E-04	3,35E-03	2,11E-03	4,01E-03	2,14E-03	2,61E-03	8,88E-04	7,21E-04	4,95E-03	8,91E-04
		9	5,05E-03	1,14E-03	1,76E-03	1,59E-03	7,05E-04	1,48E-03	1,37E-03	1,39E-03	1,68E-03	1,15E-03	1,13E-03	1,45E-03	4,00E-03	6,59E-04	1,19E-03	5,38E-03
		10	2,92E-03	1,54E-03	2,68E-03	1,60E-03	2,37E-03	8,42E-04	3,83E-03	1,11E-03	1,93E-03	2,67E-03	5,69E-04	1,22E-03	1,48E-03	1,22E-03	2,52E-03	1,23E-03
		11	8,26E-04	2,32E-03	8,04E-04	1,32E-03	1,70E-03	6,88E-04	1,17E-03	1,06E-03	1,55E-03	1,13E-03	8,01E-03	5,67E-04	5,09E-03	2,89E-03	8,40E-04	1,27E-03
		12	2,42E-03	3,20E-03	1,93E-03	1,35E-03	2,29E-03	1,24E-03	1,74E-03	1,22E-03	1,14E-03	1,01E-03	1,78E-03	4,28E-03	9,56E-04	5,81E-04	1,27E-03	1,34E-03
		13	1,54E-03	2,92E-03	1,30E-03	3,60E-03	8,32E-04	1,40E-03	6,89E-04	1,07E-03	7,52E-04	1,17E-03	1,84E-03	6,47E-04	7,20E-04	7,76E-04	4,25E-04	1,79E-03
		14	5,45E-04	2,83E-03	1,26E-03	1,39E-03	1,10E-03	7,87E-04	7,64E-04	2,12E-03	6,25E-04	1,14E-03	9,69E-04	5,23E-04	1,21E-03	1,22E-03	3,20E-03	3,94E-04
		15	9,39E-04	8,29E-03	3,87E-04	1,64E-03	2,55E-03	2,39E-03	1,14E-03	6,53E-04	1,18E-03	6,86E-04	7,00E-04	1,35E-03	1,87E-03	1,67E-03	9,61E-04	1,04E-03
		16	8,35E-04	1,25E-03	1,45E-03	1,22E-03	7,62E-04	1,31E-03	1,52E-03	8,09E-03	7,69E-04	7,63E-04	3,22E-03	1,00E-03	7,28E-04	5,84E-04	2,09E-03	1,39E-03
		17	2,18E-03	3,97E-03	1,18E-03	2,04E-03	1,64E-03	2,89E-03	9,10E-04	8,77E-04	6,66E-04	2,51E-03	1,15E-03	7,33E-04	1,66E-03	2,07E-03	1,13E-03	2,20E-03
		18	3,06E-03	5,35E-04	5,51E-04	5,51E-04	3,98E-04	2,57E-03	6,83E-04	1,61E-03	1,28E-03	6,55E-04	7,68E-04	7,53E-04	5,63E-04	6,54E-04	3,02E-03	2,37E-03
		19	1,39E-03	2,14E-03	3,75E-03	1,21E-03	7,99E-03	6,18E-04	1,12E-03	5,28E-04	1,22E-03	3,33E-03	1,06E-03	5,75E-04	8,26E-04	7,27E-04	4,92E-04	4,51E-04
		20	6,55E-04	1,42E-03	1,10E-03	2,58E-03	1,08E-03	2,79E-03	2,36E-03	2,53E-03	8,59E-04	6,59E-04	5,67E-04	1,39E-03	7,87E-03	3,52E-04	2,83E-03	9,92E-04

Πίνακας 4.19 Μέσο τετραγωνικό σφάλμα δεδομένων εκπαίδευσης με την μέθοδο trainSCG.



## ΕΠΙΛΟΓΗ ΒΕΛΤΙΣΤΟΥ ΤΝΔ ΓΙΑ ΚΑΘΕ ΜΕΘΟΔΟ ΕΚΠΑΙΔΕΥΣΗΣ

Παρακάτω παρατίθενται αρχικά για τη πρώτη τεχνική επιλογής βέλτιστων ΤΝΔ, τρεις πίνακες (ένας για κάθε μέθοδο εκπαίδευσης) που αναφέρονται στα ΤΝΔ με τον μικρότερο αριθμό δεδομένων με σφάλμα >10% (1 δεδομένο > 10% για τις δομές που εκπαιδεύτηκαν με την μέθοδο trainGDX και 0 δεδομένα >10% για τις μεθόδους trainLM και trainSCG). Στη συνέχεια παρατίθενται άλλοι τρεις αντίστοιχοι πίνακες για την δεύτερη τεχνική όπως αναφέραμε παραπάνω, ελέγχοντας δηλαδή πρώτα τα mse ελέγχου και εκπαίδευσης και στη συνέχεια τον αριθμό δεδομένων με σφάλμα >10%. Σχετικά με τους χρωματισμούς κάποιον κελιών ισχύει το εξής:

Για την πρώτη τεχνική: το απαλό πράσινο δείχνει τις δύο δομές με το μικρότερο μέσο τετραγωνικό σφάλμα ελέγχου και το έντονο πράσινο δείχνει ποιά δομή από αυτές τις δύο είχε το μικρότερο μέσο τετραγωνικό σφάλμα εκπαίδευσης (που μας οδηγεί στην επιλογή του βέλτιστου ΤΝΔ).

Για την δεύτερη τεχνική: το απαλό πράσινο δείχνει τις δύο δομές με το μικρότερο μέσο τετραγωνικό σφάλμα εκπαίδευσης και το έντονο πράσινο δείχνει ποιά δομή από αυτές τις δύο είχε το μικρότερο αριθμό δεδομένων με σφάλμα >10% (που μας οδηγεί στην επιλογή του βέλτιστου ΤΝΔ).

### Αποτελέσματα με χρήση της πρώτης τεχνικής:

trainGDX				
#	Δομή	>10%	mse Ελέγχου	mse Εκπαίδευσης
1	10x18	1	7,64 E-04	9,55 E-04
2	12x14	1	3,86 E-04	3,69 E-04
3	13x7	1	3,55 E-04	6,13 E-04
4	13x18	1	2,66 E-04	7,60 E-04
5	13x20	1	4,38 E-04	5,54 E-04
6	14x13	1	3,74 E-04	6,62 E-04
7	15x16	1	2,68 E-04	5,08 E-04
8	17x5	1	1,50 E-04	6,96 E-04
9	18x13	1	5,30 E-04	5,72 E-04
10	19x10	1	2,29 E-04	4,65 E-04
11	20x14	1	3,06 E-04	5,01 E-04

**Βέλτιστο**

trainLM				
#	Δομή	>10%	mse Ελέγχου	mse Εκπαίδευσης
1	6x12	0	2,83 E-04	9,01 E-04
2	6x13	0	3,05 E-04	2,48 E-04
3	7x13	0	2,53 E-04	3,41 E-04
4	11x5	0	1,20 E-04	2,60 E-04
5	11x16	0	5,12 E-04	3,44 E-04
6	15x6	0	3,63 E-04	5,16 E-04
7	15x7	0	3,15 E-04	4,98 E-04

**Βέλτιστο**

trainSCG					
#	Δομή	>10%	mse Ελέγχου	mse Εκπαίδευσης	
1	12x18	0	2,96 E-04	5,81 E-04	
2	16x18	0	3,64 E-04	5,84 E-04	
3	18x9	0	2,83 E-04	3,98 E-04	<b>Βέλτιστο</b>
4	19x19	0	5,05 E-04	4,92 E-04	
5	20x18	0	4,55 E-04	3,52 E-04	

Αποτελέσματα με χρήση της δεύτερης τεχνικής:

trainGDX					
#	Δομή	mse Ελέγχου	mse Εκπαίδευσης	>10%	
1	7x17	6,01 E-05	2,15 E-03	7	
2	17x9	6,45 E-05	9,21 E-04	6	<b>Βέλτιστο</b>
3	15x5	8,60 E-05	1,61 E-03	5	
4	13x9	9,25 E-05	2,08 E-03	10	
5	13x15	9,39 E-05	1,78 E-03	9	
6	15x13	9,89 E-05	2,75 E-03	10	

trainLM					
#	Δομή	mse Ελέγχου	mse Εκπαίδευσης	>10%	
1	5x7	3,73 E-05	7,11 E-04	3	<b>Βέλτιστο</b>
2	5x8	5,85 E-05	9,08 E-04	3	
3	7x19	6,29 E-05	8,15 E-04	3	
4	6x5	7,77 E-05	7,53 E-04	4	
5	16x6	8,88 E-05	2,58 E-03	12	
6	18x19	9,23 E-05	1,11 E-03	5	
7	5x14	9,34 E-05	9,02 E-04	5	

trainSCG					
#	Δομή	mse Ελέγχου	mse Εκπαίδευσης	>10%	
1	9x6	4,54 E-05	1,14 E-03	1	
2	13x7	6,33 E-05	1,30 E-03	4	
3	18x15	6,40 E-05	7,68 E-04	3	<b>Βέλτιστο</b>
4	11x13	7,89 E-05	1,55 E-03	9	
5	15x7	8,32 E-05	3,87 E-04	2	
6	8x14	8,36 E-05	4,01 E-03	18	
7	15x6	9,02 E-05	8,29 E-03	19	

**Πίνακας 4.20** Επιλογή βέλτιστων ΤΝΔ για κάθε μέθοδο εκπαίδευσης.

Επειδή θέλουμε να επιλέξουμε ως τελικό βέλτιστο ΤΝΔ ένα δίκτυο που σίγουρα γενικεύει καλά ενώ ταυτόχρονα δίνει όσο το δυνατόν περισσότερο "ομαλά" αποτελέσματα (υπό την έννοια ότι το σφάλμα είναι κατανομημένο σε όλα τα δεδομένα και δεν υπάρχουν ακραίες διακυμάνσεις του), κρατάμε τα τρία ΤΝΔ που προέκυψαν από την πρώτη τεχνική.

Στη συνέχεια παρατίθενται για κάθε ένα από τα τρία βέλτιστα ΤΝΔ τα διαγράμματα:

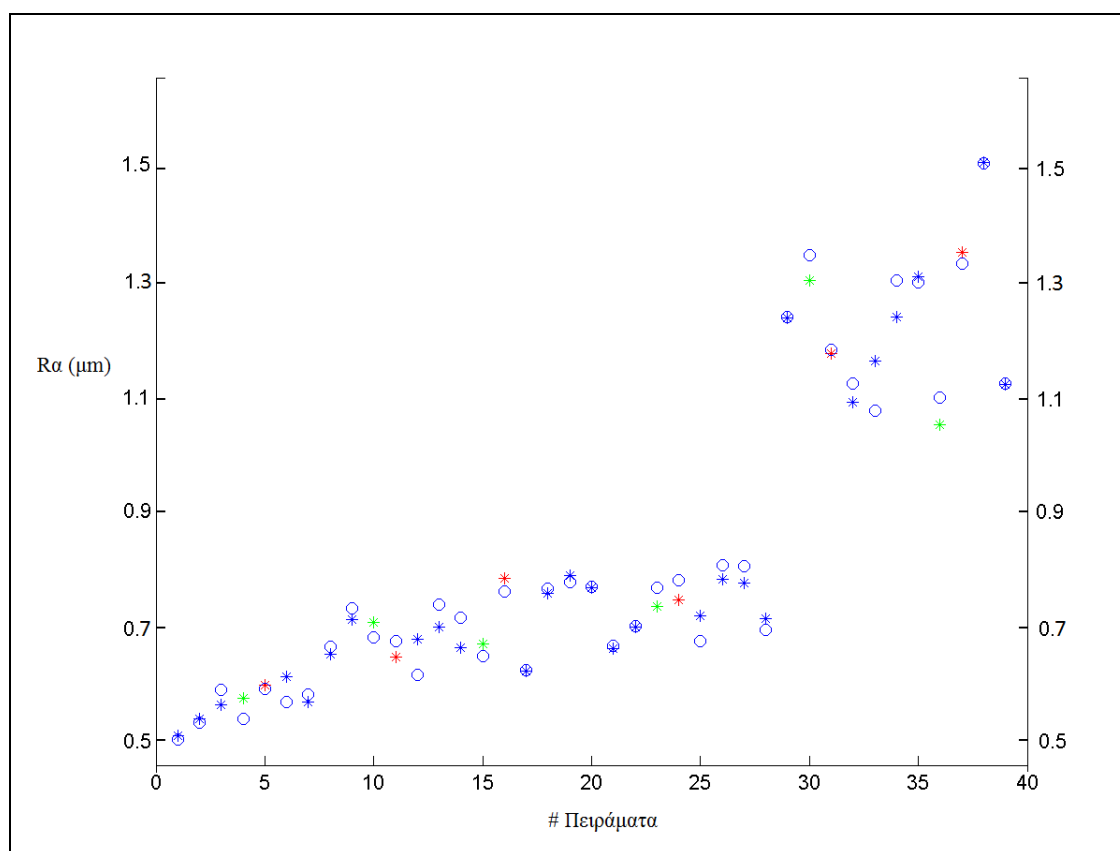
- Τα πειραματικά δεδομένα και οι αντίστοιχες τιμές επιφανειακής τραχύτητας από το ΤΝΔ
- Regression των set: train, validation, test
- Απόδοσης του δικτύου κατά την εκπαίδευση

Στα διαγράμματα αυτά, το μπλε χρώμα αναφέρεται σε δεδομένα εκπαίδευσης, το πράσινο σε δεδομένα validation και το κόκκινο σε δεδομένα ελέγχου.

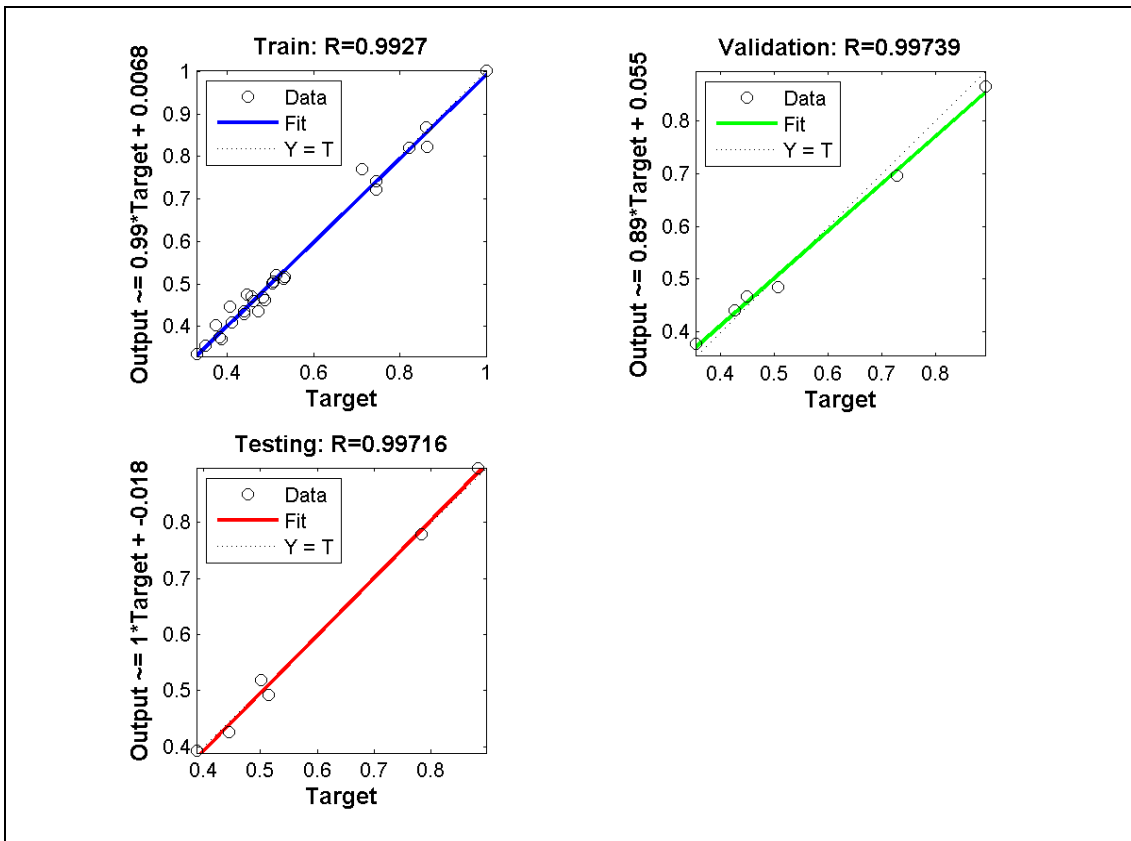
Στο διάγραμμα επιφανειακής τραχύτητας, τα πειραματικά δεδομένα που αποτελούν και στόχο του ΤΝΔ παρουσιάζονται ως "κυκλάκια" ενώ οι τιμές που δίνει το ΤΝΔ ως "αστερίσκοι".

Στο διάγραμμα regression υπάρχουν τρία ξεχωριστά υποδιαγράμματα τα οποία αναφέρονται στο train set (με μπλε χρώμα), στο validation set (με πράσινο χρώμα) και στο test set (με κόκκινο χρώμα). Ο κατακόρυφος άξονας αναφέρεται στις τιμές που δίνει το δίκτυο (κανονικοποιημένες) ενώ στον οριζόντιο οι (κανονικοποιημένες) τιμές που δίνει το ΤΝΔ. Στόχος είναι η χρωματιστή γραμμή που προκύπτει από τα αποτελέσματα να περνά όσο το δυνατόν πλησιέστερα της διαγωνίου. (προσοχή: στα διαγράμματα regression τα πειράματα δεν παρουσιάζονται στους άξονες με την σειρά που έγιναν, αλλά με σειρά ώστε η επιφανειακή τραχύτητα να αυξάνεται καθώς κινούμαστε στον κάθε άξονα).

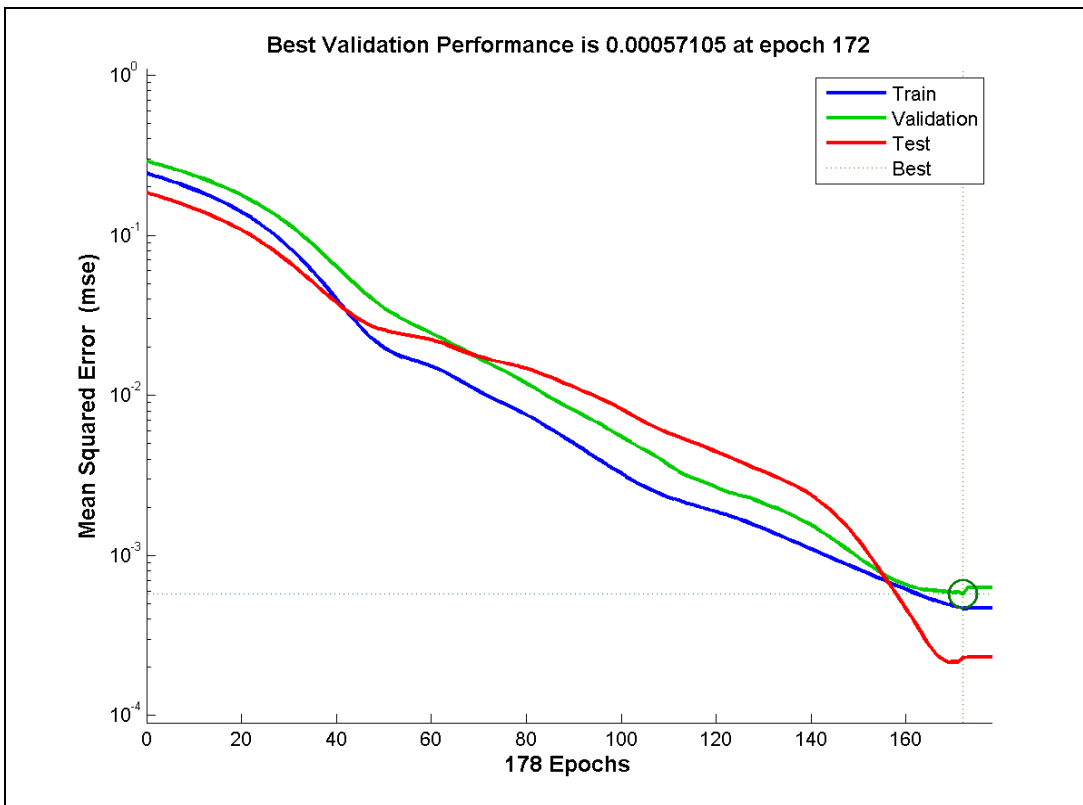
Στο διάγραμμα απόδοσης φαίνονται τα τρία μέσα τετραγωνικά σφάλματα (train, validation και test) καθώς προχωρούν οι εποχές (οριζόντιος άξονας) κατά την εκπαίδευση. Με κύκλο επισημαίνεται η εποχή που επιλέχθηκε με βάση το κριτήριο του σφάλματος του validation set η οποία αναφέρεται και στο πάνω μέρος του διαγράμματος. Παρατηρούμε από τον οριζόντιο άξονα ότι οι εποχές των διαγραμμάτων φθάνουν έως και 6 επιπλέον εποχές, μετά την βέλτιστη εποχή, οι οποίες ενεργοποίησαν το κριτήριο.



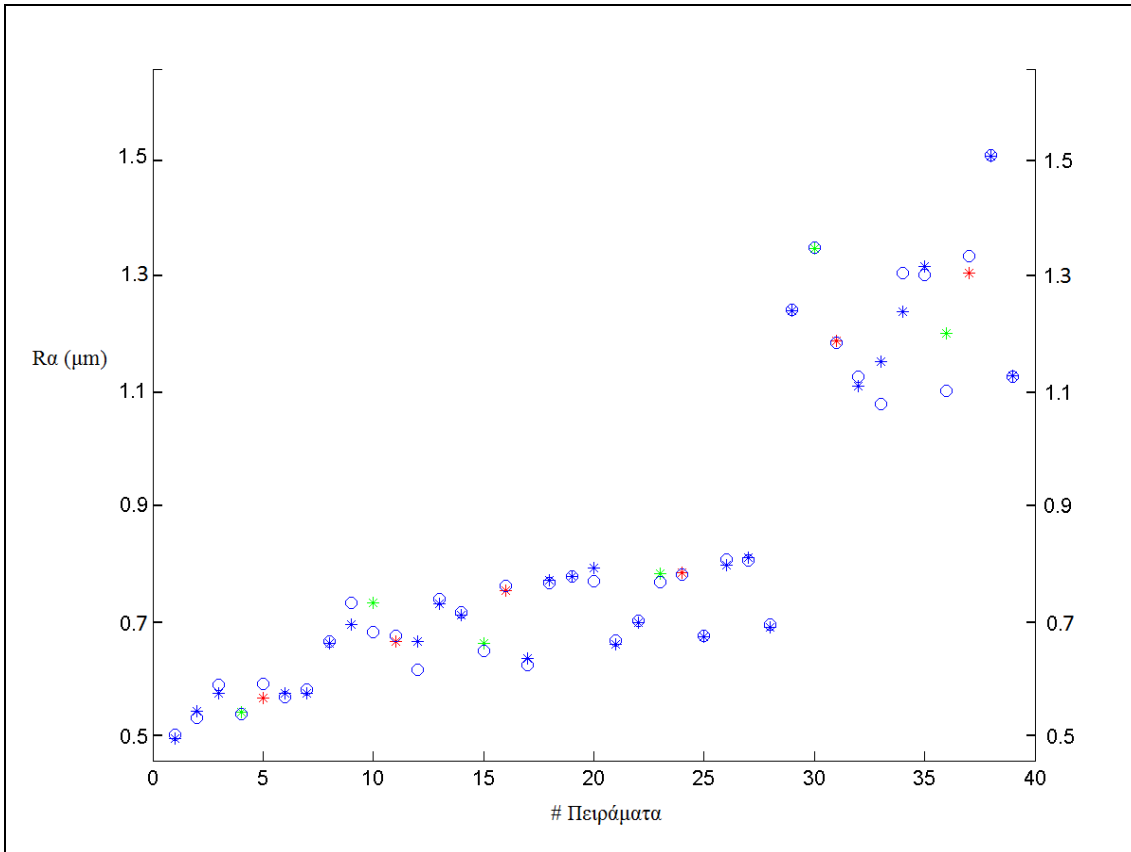
**Διάγραμμα 4.4** Τα πειραματικά δεδομένα και οι αντίστοιχες τιμές επιφανειακής τραχύτητας του βέλτιστου ΤΝΔ της μεθόδου εκπαίδευσης trainGDX.



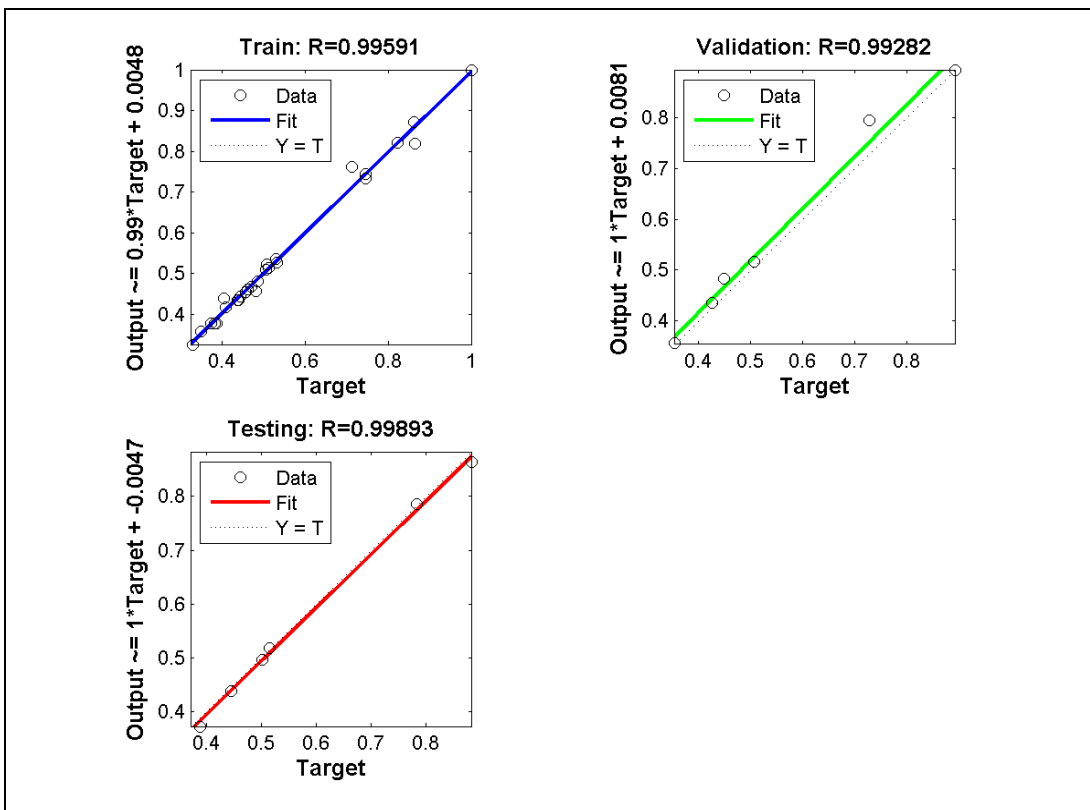
**Διάγραμμα 4.5** Τα διαγράμματα Regression των τριών σετ δεδομένων για το βέλτιστο ΤΝΔ της μεθόδου εκπαίδευσης trainGDX.



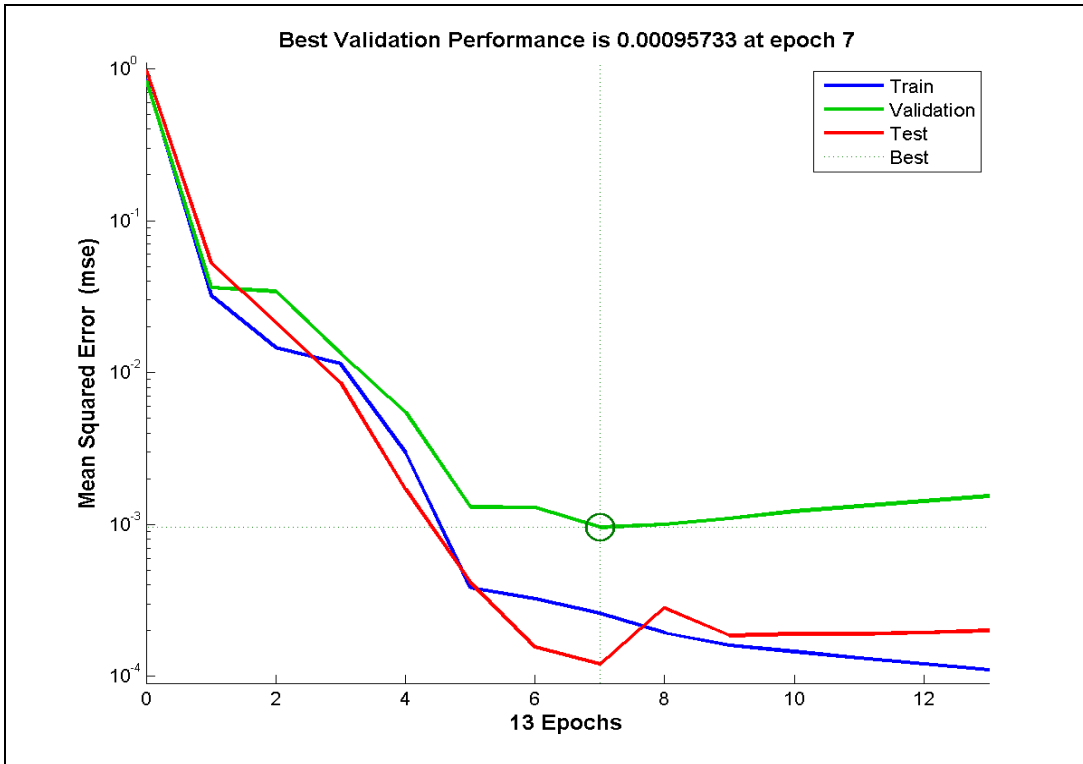
**Διάγραμμα 4.6** Το διάγραμμα απόδοσης του βέλτιστου ΤΝΔ με τη μέθοδο trainGDX κατά την εκπαίδευσή του.



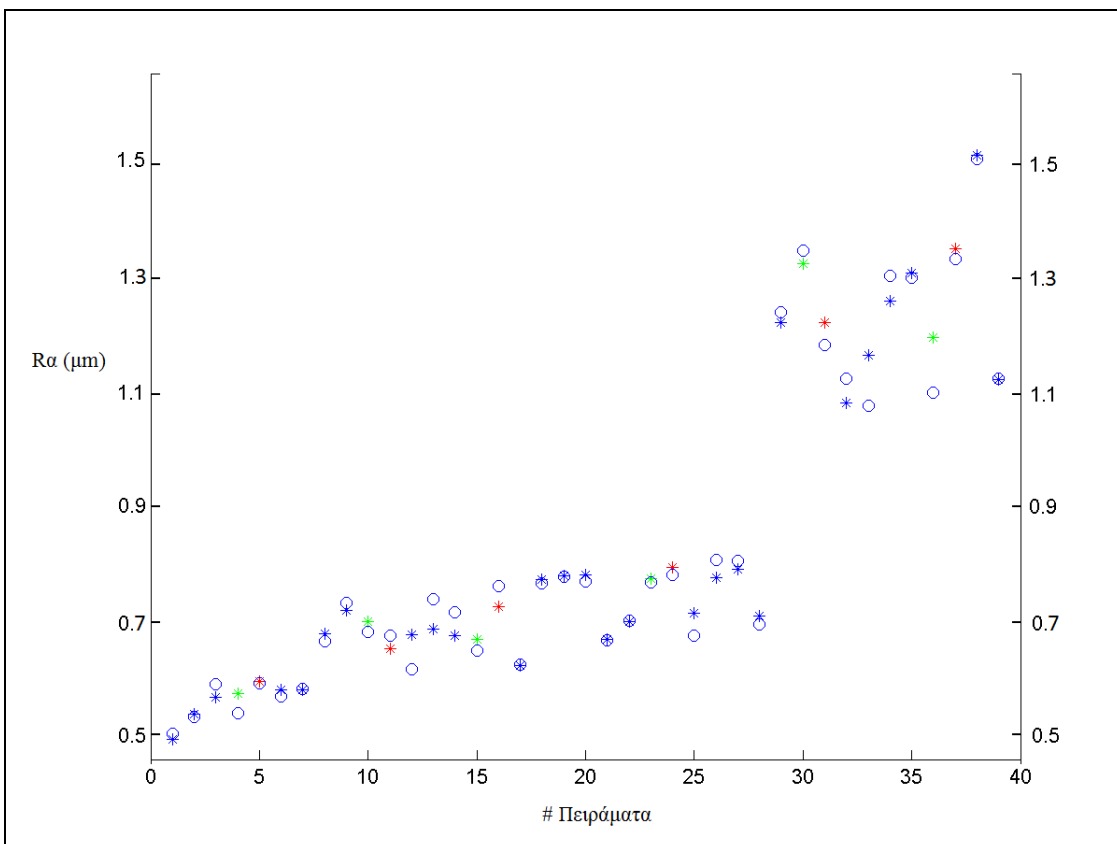
**Διάγραμμα 4.7** Τα πειραματικά δεδομένα και οι αντίστοιχες τιμές επιφανειακής τραχύτητας του βέλτιστου ΤΝΔ της μεθόδου εκπαίδευσης trainLM.



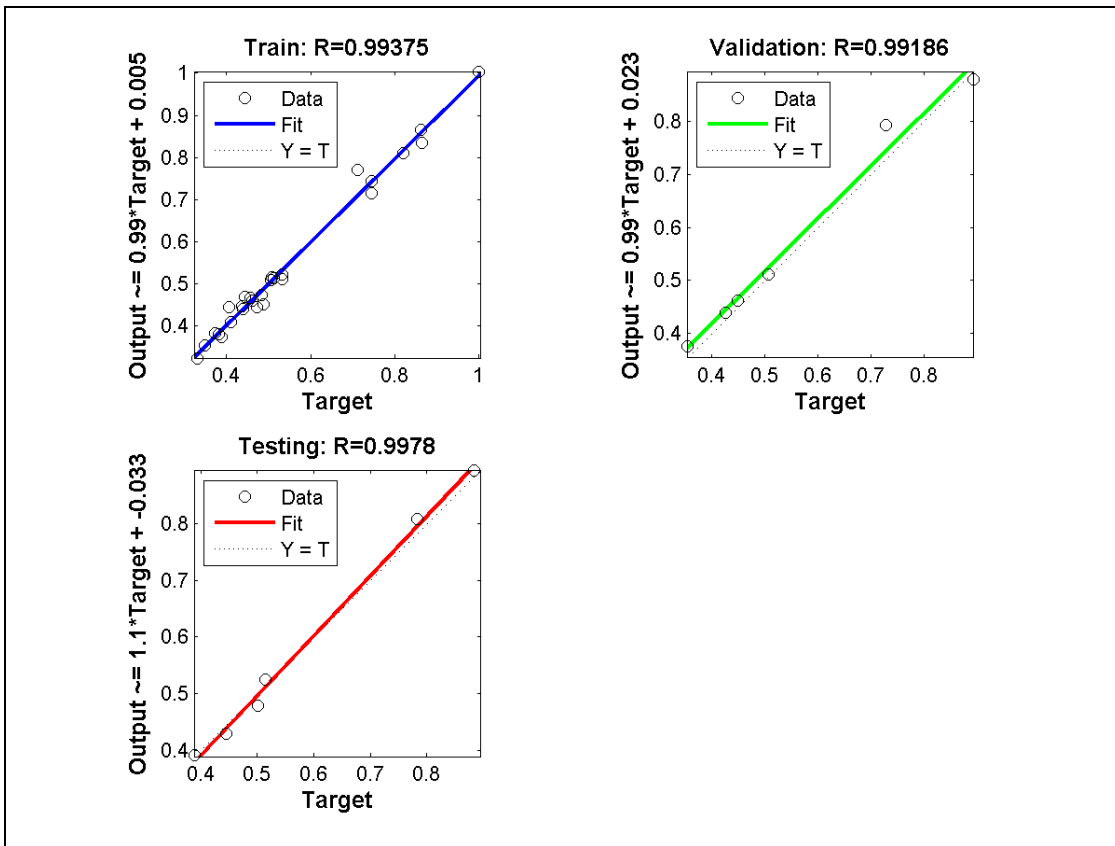
**Διάγραμμα 4.8** Τα διαγράμματα Regression των τριών σετ δεδομένων για το βέλτιστο ΤΝΔ της μεθόδου trainLM.



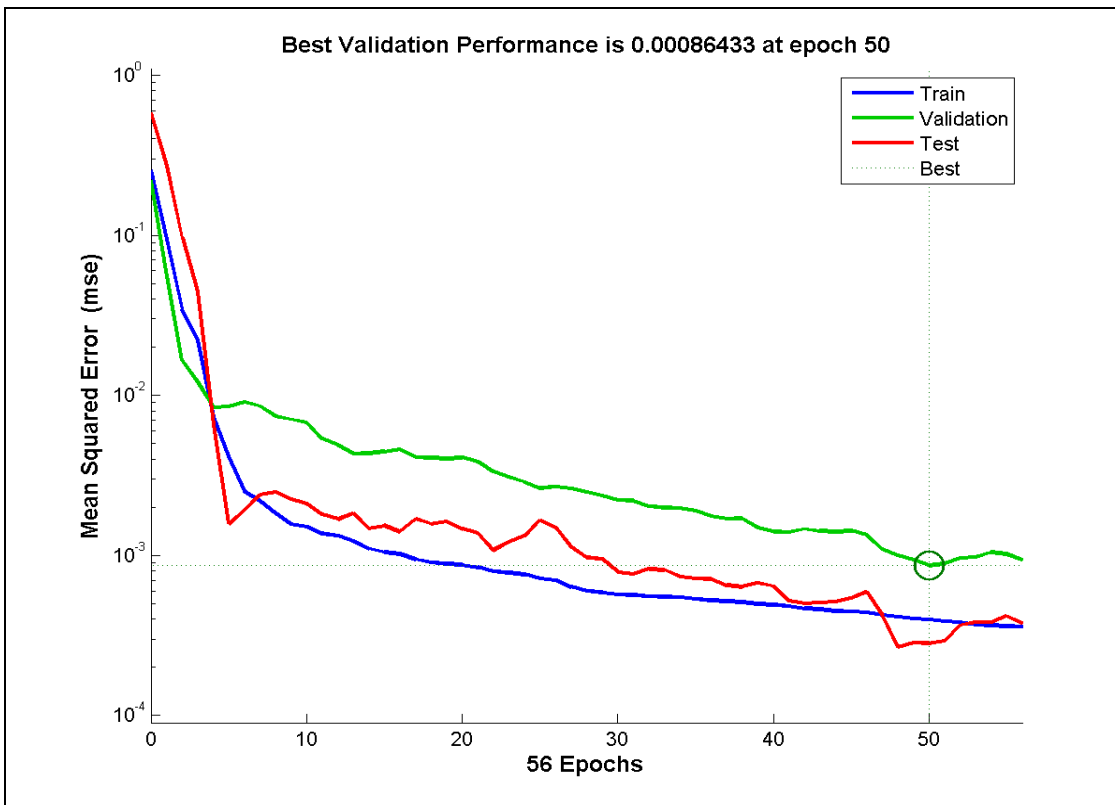
**Διάγραμμα 4.9** Το διάγραμμα απόδοσης του βέλτιστου ΤΝΔ με τη μέθοδο trainLM κατά την εκπαίδευσή του.



**Διάγραμμα 4.10** Τα πειραματικά δεδομένα και οι αντίστοιχες τιμές επιφανειακής τραχύτητας του βέλτιστου ΤΝΔ της μεθόδου εκπαίδευσής trainSCG.



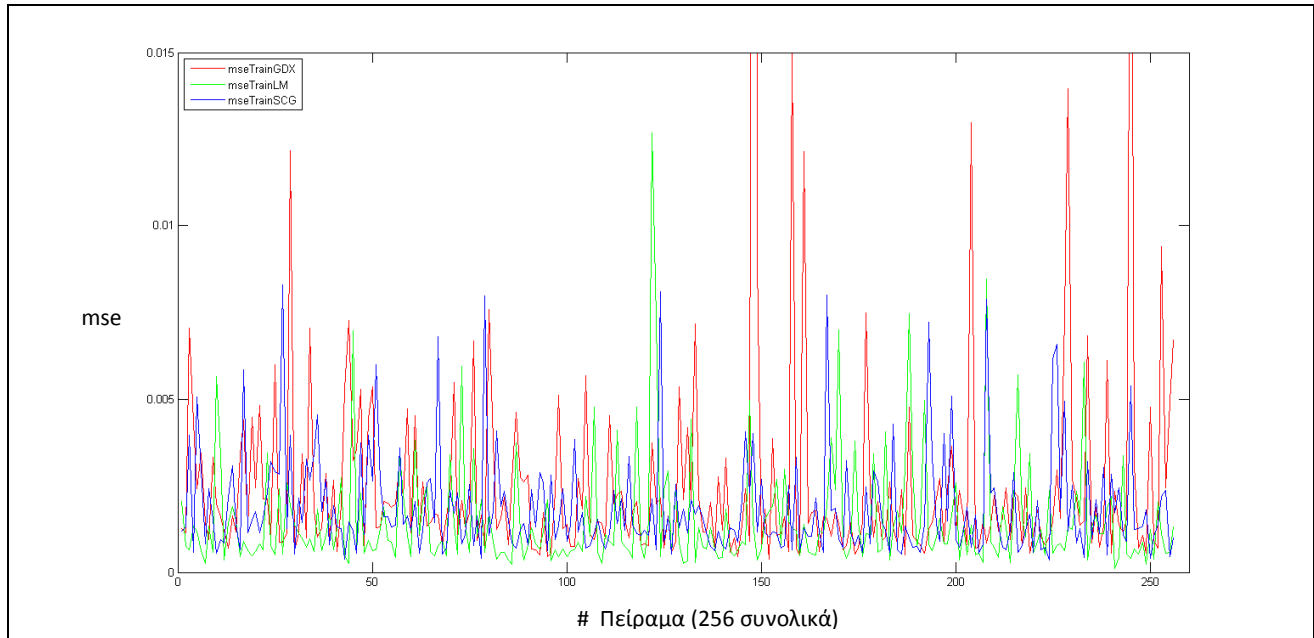
**Διάγραμμα 4.11** Τα διαγράμματα Regression των τριών σετ δεδομένων για το βέλτιστο ΤΝΔ της μεθόδου trainSCG.



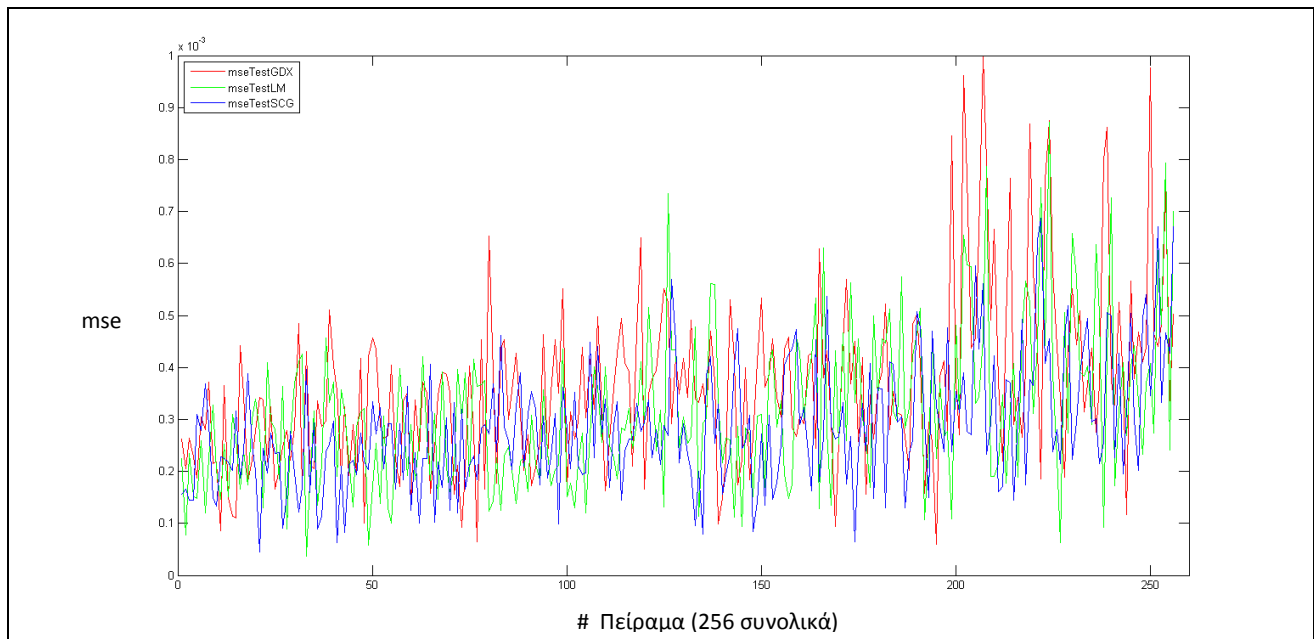
**Διάγραμμα 4.12** Το διάγραμμα απόδοσης του βέλτιστου ΤΝΔ με τη μέθοδο trainSCG κατά την εκπαίδευσή του.

## ΠΑΡΑΤΗΡΗΣΕΙΣ ΑΠΟ ΔΙΑΓΡΑΜΜΑΤΑ ΓΙΑ ΤΙΣ ΜΕΘΟΔΟΥΣ ΕΚΠΑΙΔΕΥΣΗΣ

Παρακάτω παρατίθενται δύο διαγράμματα ώστε να βγάλουμε ορισμένα συμπεράσματα για τις τρεις μεθόδους εκπαίδευσης που χρησιμοποιήθηκαν. Στον οριζόντιο άξονα βρίσκονται οι 256 διαφορετικές δομές ΤΝΔ με τη σειρά που παρατίθενται και στους πίνακες συνολικών αποτελεσμάτων. Στον κατακόρυφο άξονα βρίσκονται οι τιμές του MSE.



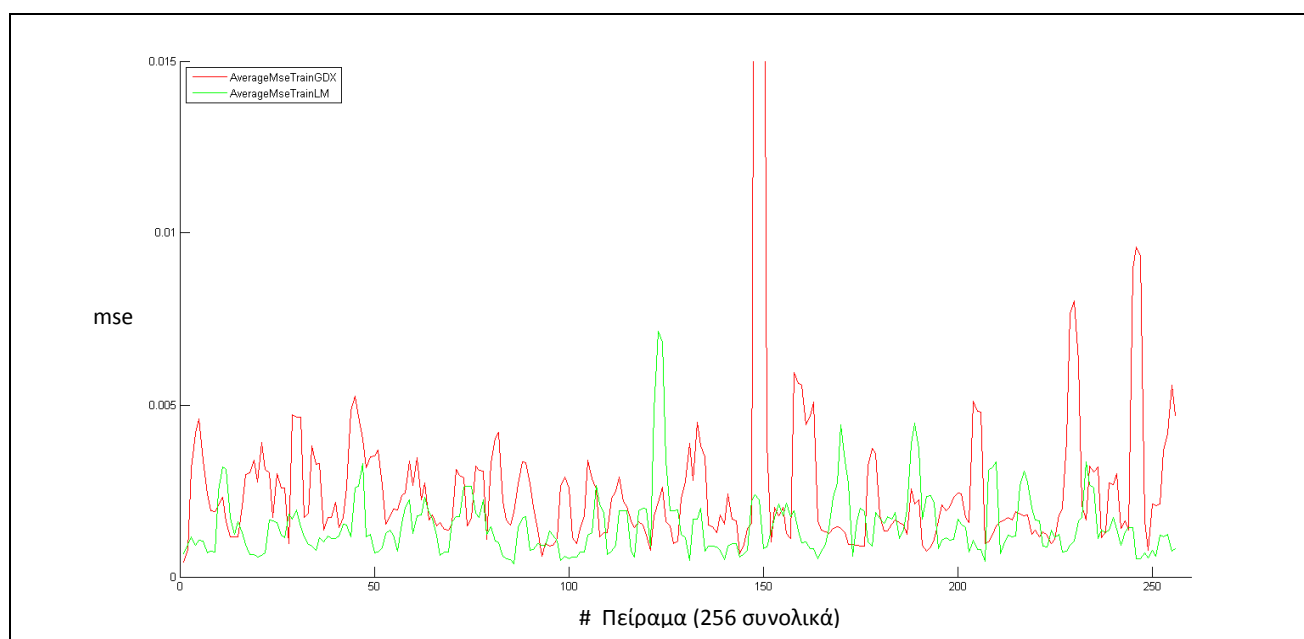
**Διάγραμμα 4.13** Το μέσο τετραγωνικό σφάλμα των δεδομένων εκπαίδευσης για κάθε μέθοδο.



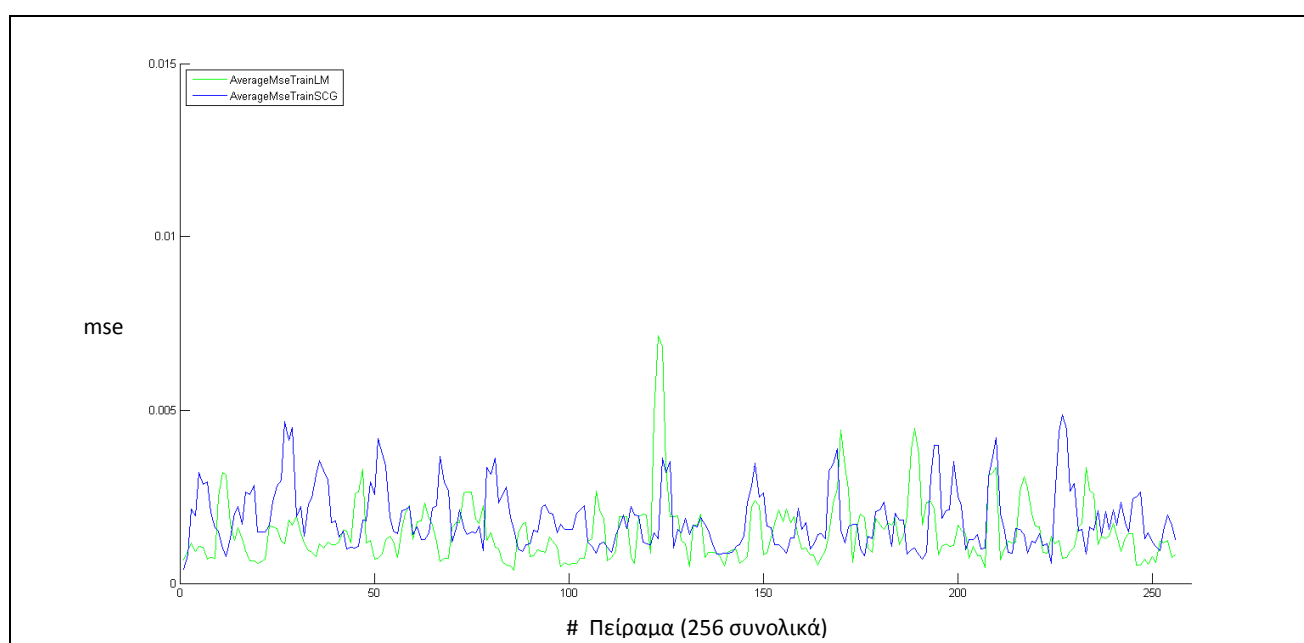
**Διάγραμμα 4.14** Το μέσο τετραγωνικό σφάλμα των δεδομένων ελέγχου για κάθε μέθοδο.



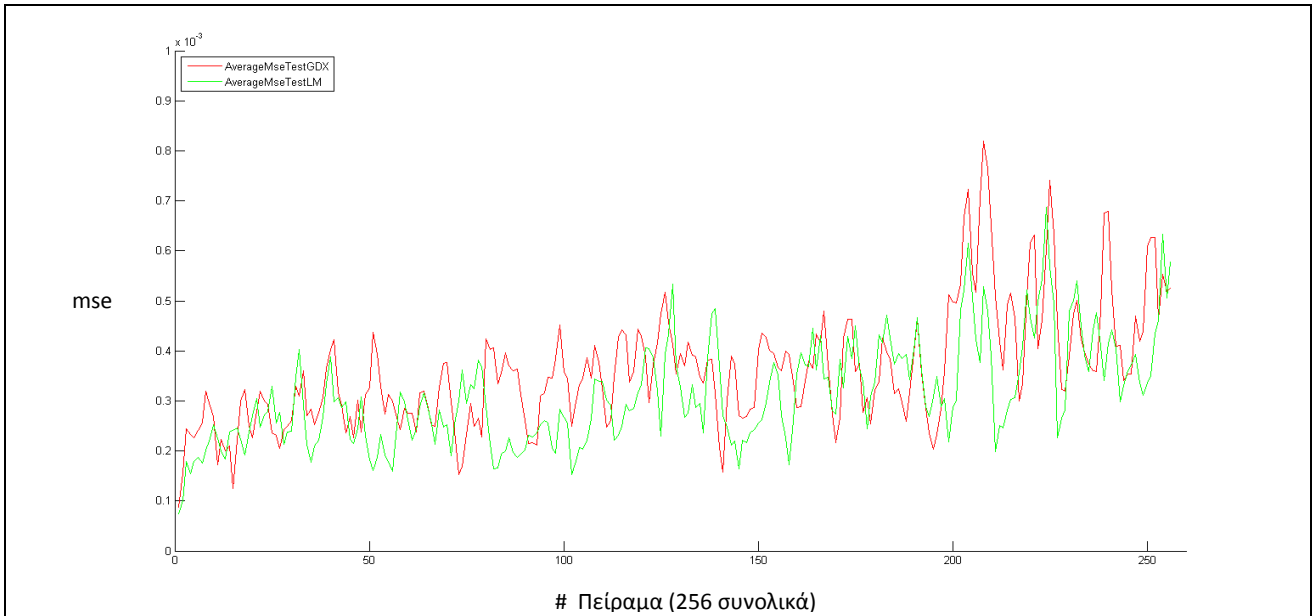
Επειδή με τα παραπάνω διαγράμματα είναι δύσκολο να εξάγουμε συμπεράσματα, παρακάτω παραθέτουμε δύο συγκριτικά διαγράμματα συγκρίνοντας τη μέθοδο trainGDX με την trainLM και τη μέθοδο trainLM με την trainSCG, αντίστοιχα, για το μέσο τετραγωνικό σφάλμα των δεδομένων εκπαίδευσης. Στη συνέχεια παραθέτουμε δύο ακόμα αντίστοιχα διαγράμματα για το μέσο τετραγωνικό σφάλμα των δεδομένων ελέγχου. Να σημειωθεί ότι στα συγκεκριμένα διαγράμματα οι τιμές MSE δεν αντιστοιχούν στις πραγματικές τιμές όπως στα δύο προηγούμενα διαγράμματα αλλά φιλτραρίστηκαν και ομαλοποιήθηκαν για να είναι περισσότερο εμφανείς οι διαφορές μεταξύ των μεθόδων. Όμοια με προηγουμένως, στον οριζόντιο άξονα βρίσκονται οι 256 διαφορετικές δομές TND με τη σειρά που παρατίθενται και στους πίνακες συνολικών αποτελεσμάτων, ενώ κατακόρυφο άξονα βρίσκονται οι τιμές του MSE.



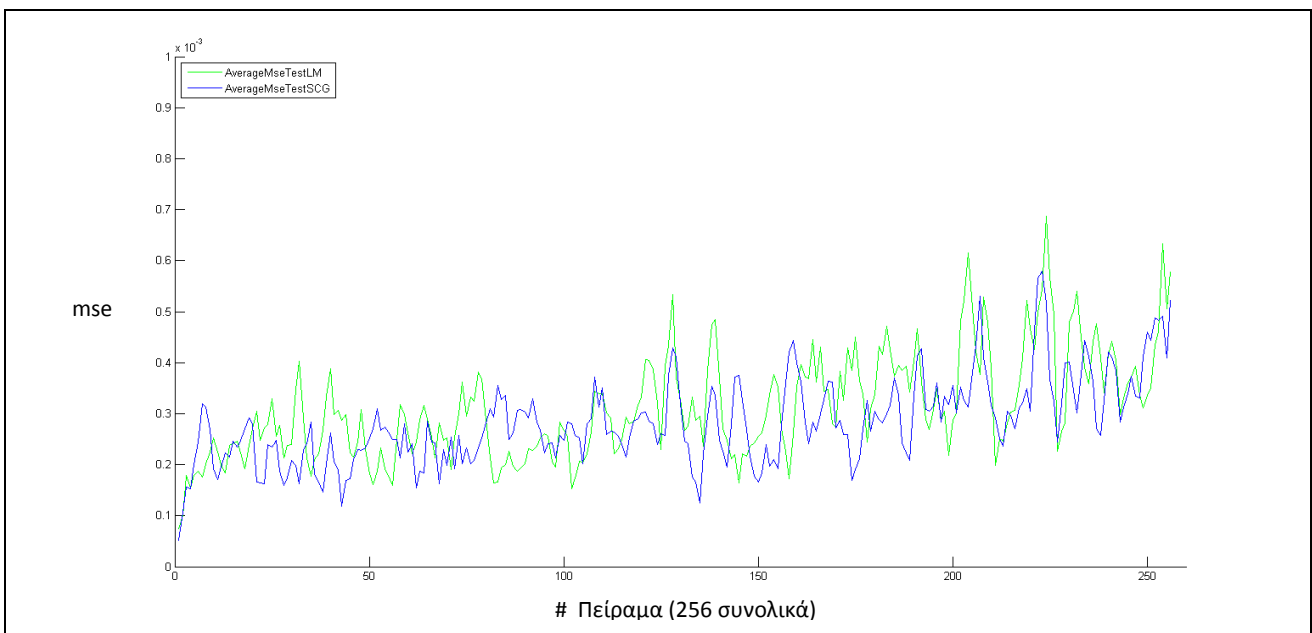
**Διάγραμμα 4.15** Συγκριτικό διάγραμμα για τις μεθόδους trainGDX και trainLM για το μέσο τετραγωνικό σφάλμα των δεδομένων εκπαίδευσης.



**Διάγραμμα 4.16** Συγκριτικό διάγραμμα για τις μεθόδους trainLM και trainSCG για το μέσο τετραγωνικό σφάλμα των δεδομένων εκπαίδευσης.



**Διάγραμμα 4.17** Συγκριτικό διάγραμμα για τις μεθόδους trainGDX και trainLM για το μέσο τετραγωνικό σφάλμα των δεδομένων ελέγχου.

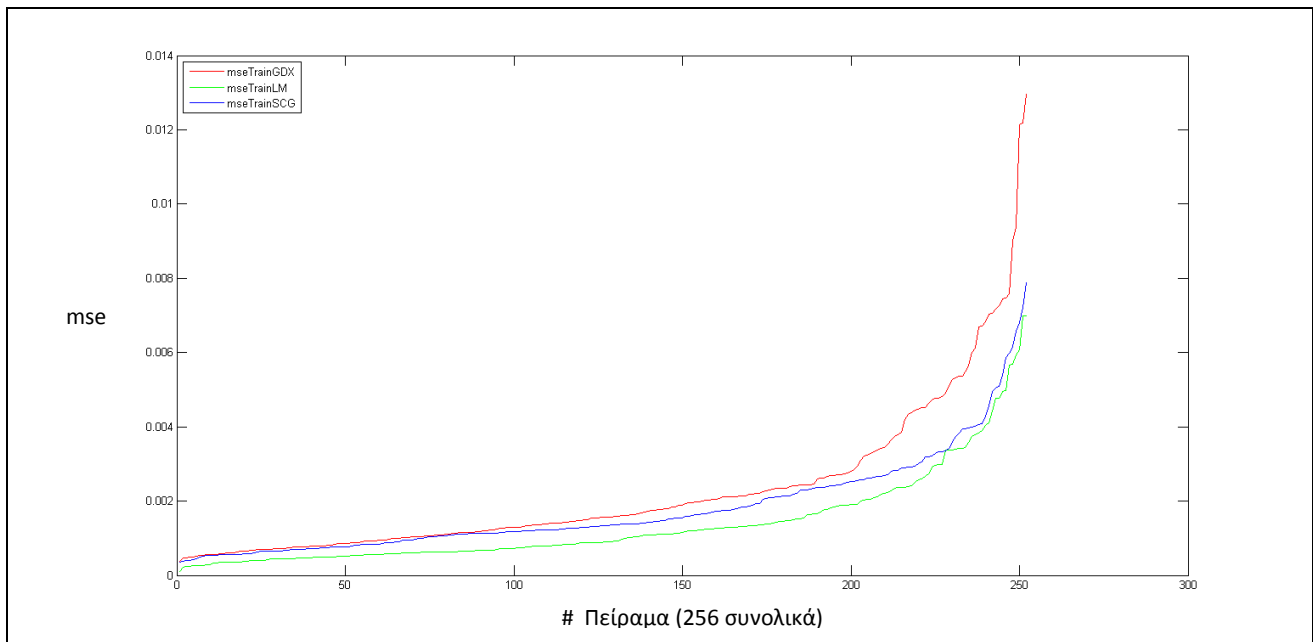


**Διάγραμμα 4.18** Συγκριτικό διάγραμμα για τις μεθόδους trainLM και trainSCG για το μέσο τετραγωνικό σφάλμα των δεδομένων ελέγχου.

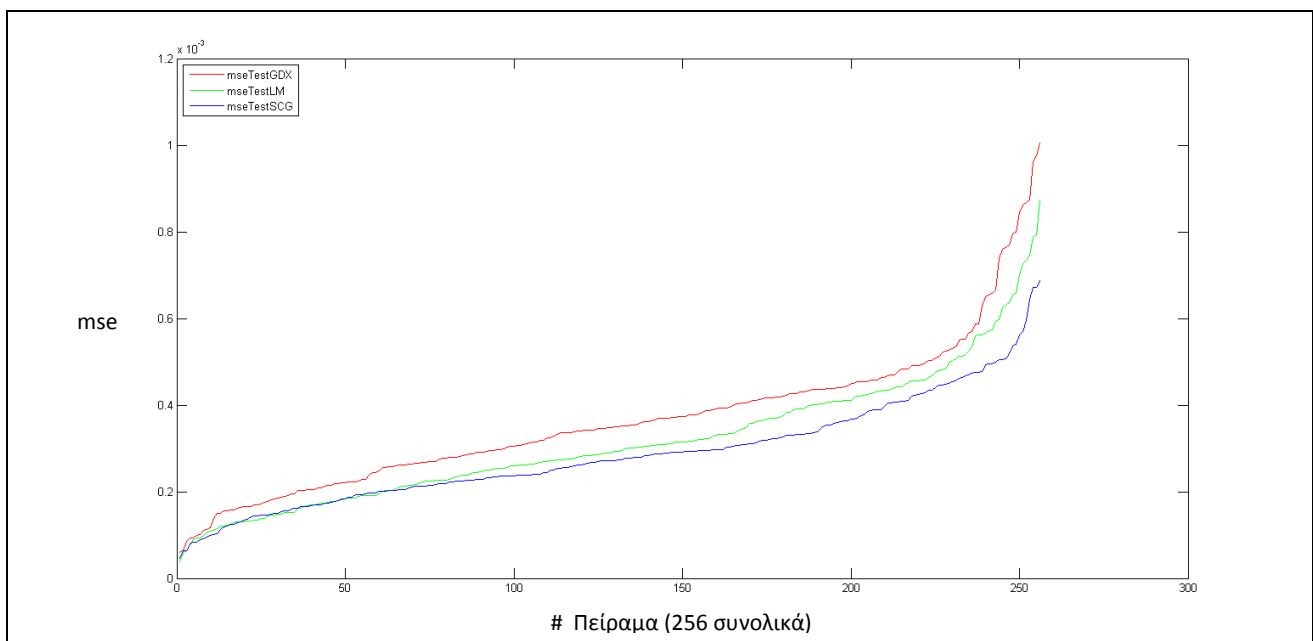
Παρατηρώντας τα παραπάνω διαγράμματα μπορούμε να καταλήξουμε στα εξής συμπεράσματα:

- Όσον αφορά το μέσο τετραγωνικό σφάλμα εκπαίδευσης, φαίνεται η μέθοδος trainGDX να έχει δώσει τα χειρότερα αποτελέσματα από τις τρεις μεθόδους. Ανάμεσα στην trainLM και την trainSCG φαίνεται η μέθοδος trainLM να έχει οριακά μικρότερα σφάλματα.
- Όσον αφορά το μέσο τετραγωνικό σφάλμα ελέγχου, φαίνεται η μέθοδος trainGDX να έχει οριακά χειρότερα αποτελέσματα από τη μέθοδο trainLM, ενώ από το συγκριτικό διάγραμμα 4.18 των μεθόδων trainLM και trainSCG δεν μπορούμε να βγάλουμε κάποιο συμπέρασμα.

Επειδή με τον οπτικό έλεγχο των παραπάνω διαγραμμάτων οι διαφορές είναι δυσδιάκριτες, παραθέτουμε στη συνέχεια δύο διαγράμματα για τα δεδομένα εκπαίδευσης και ελέγχου αντίστοιχα όπου στον οριζόντιο άξονα έχουμε τις 256 διαφορετικές δομές των ΤΝΔ με τέτοια σειρά ώστε τα πραγματικά MSEs να έχουν αύξουσα τιμή.



**Διάγραμμα 4.19** Συγκριτικό διάγραμμα των τριών μεθόδων για τα δεδομένα εκπαίδευσης (MSE εκπαίδευσης με αύξουσα τιμή).



**Διάγραμμα 4.20** Συγκριτικό διάγραμμα των τριών μεθόδων για τα δεδομένα ελέγχου (MSE ελέγχου με αύξουσα τιμή).

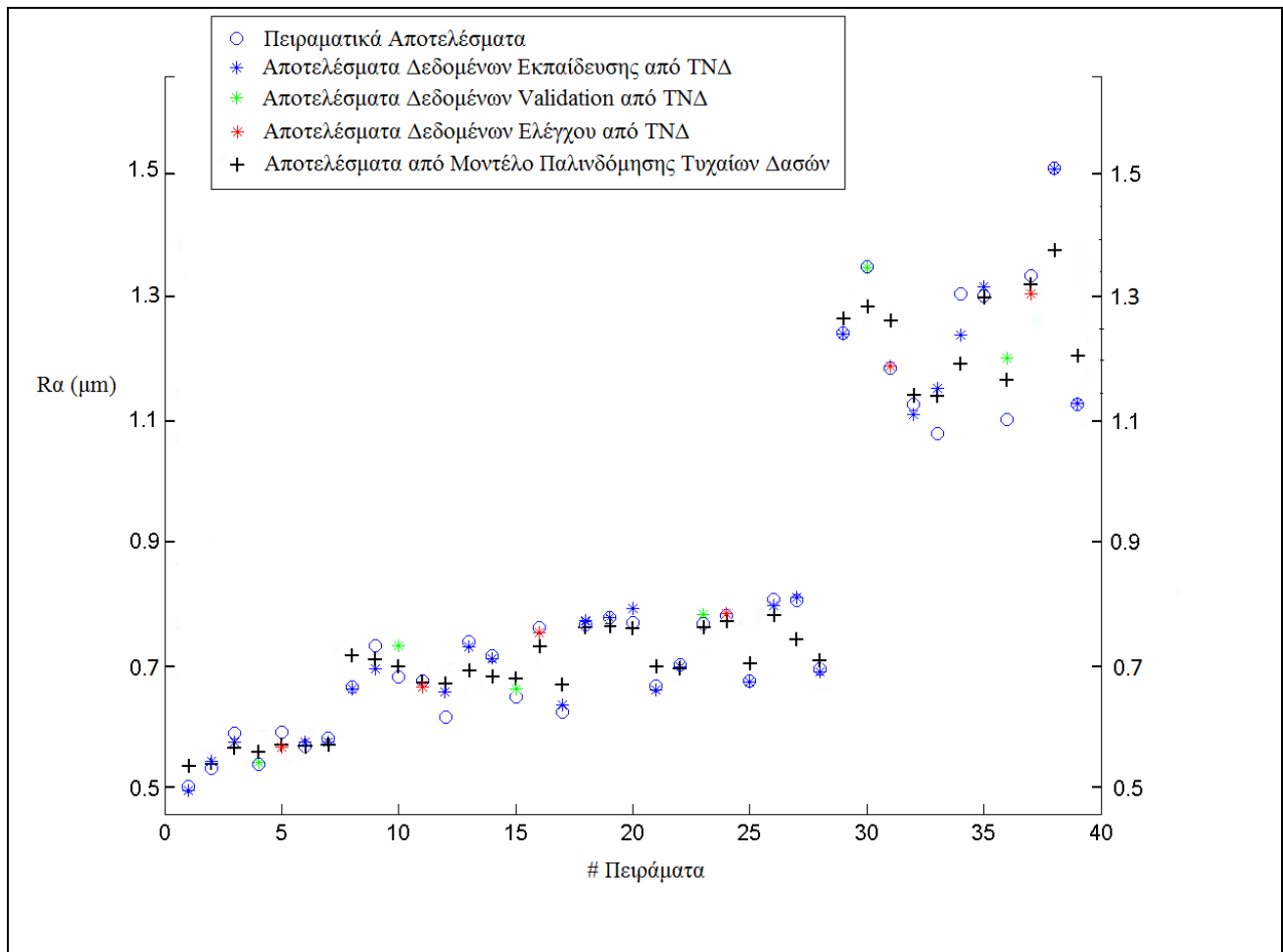
Από τα παραπάνω διαγράμματα μπορούμε να επαληθεύσουμε τα συμπεράσματα που εξάγαμε προηγουμένως καθώς επίσης διαπιστώνουμε και ότι η μέθοδος trainSCG έδωσε μικρότερα σφάλματα ελέγχου από τη μέθοδο trainLM.

Όσον αφορά τον ρόλο που παίζει το μέγεθος του δικτύου (ο αριθμός των νευρώνων στα κρυφά επίπεδα) στην απόδοσή του, μπορούμε να παρατηρήσουμε ότι καθώς κινούμαστε προς τα δεξιά (δηλαδή προς την αύξηση του αριθμού νευρώνων στο πρώτο κρυφό επίπεδο), φαίνεται να αυξάνεται ελαφρώς το σφάλμα ελέγχου για όλες τις μεθόδους (προσοχή, αυτό το συμπέρασμα δεν εξάγεται από τα διαγράμματα 4.19 και 4.20 όπου οι δομές στον οριζόντιο άξονα είναι ανακατεμένες). Για το σφάλμα εκπαίδευσης δεν παρατηρείται κάποιο αντίστοιχο συμπέρασμα.

## ΣΥΓΚΡΙΣΗ ΒΕΛΤΙΣΤΟΥ ΤΝΔ ΜΕ ΤΟ ΜΟΝΤΕΛΟ ΠΑΛΙΝΔΡΟΜΗΣΗΣ ΤΥΧΑΙΩΝ ΔΑΣΩΝ

Από τα τρία βέλτιστα ΤΝΔ θα επιλέξουμε για σύγκριση με το μοντέλο παλινδρόμησης τυχαίων δασών, εκείνο το οποίο έχει το μικρότερο μέσο τετραγωνικό σφάλμα εκπαίδευσης και ελέγχου. Το δίκτυο αυτό είναι το βέλτιστο ΤΝΔ της μεθόδου trainLM με δομή 3-11-5-1.

Παρακάτω παρατίθεται ένα διάγραμμα όπου φαίνονται οι αποκρίσεις του μοντέλου παλινδρόμησης τυχαίων δασών και οι αποκρίσεις του ΤΝΔ της μεθόδου trainLM με δομή 3-11-5-1 σε σύγκριση με τις πειραματικές μετρήσεις:



**Διάγραμμα 4.21** Συγκριτικό διάγραμμα μοντέλου παλινδρομησης τυχαίων δασών και βέλτιστου ΤΝΔ με δομή 3-11-5-1.

Από το παραπάνω διάγραμμα είναι εμφανές ότι το ΤΝΔ που επιλέξαμε ως βέλτιστο δίνει σαφώς καλύτερα αποτελέσματα από το μοντέλο παλινδρόμησης με τυχαία δάση το οποίο, από τα συνολικά 39 πειράματα, δίνει οριακά καλύτερο αποτέλεσμα μόνο στα πειράματα 5, 9, 10, 11, 20, 23, 33, 35, 36 και 37. Επίσης, το μοντέλο παλινδρόμησης τυχαίων δασών παρουσιάζει σημαντικά μεγαλύτερο σφάλμα σε σχέση με τα αποτελέσματα του ΤΝΔ σε αρκετά δεδομένα (όπως στα πειράματα 1, 8, 13, 27, 30, 31, 38 και 39).

#### 4.4 Συμπεράσματα

Στα πλαίσια της παρούσας μεταπτυχιακής εργασίας επιχειρήσαμε να βρούμε το βέλτιστο τεχνητό νευρωνικό δίκτυο, το οποίο θα μπορούσε να κάνει πρόβλεψη της τιμής της επιφανειακής τραχύτητας σε τórνευση του χάλυβα AISI 4340 (69HRC). Για να καταλήξουμε σε ένα τέτοιο ΤΝΔ ακολουθήσαμε την παρακάτω διαδικασία:

1. Δοκιμάσαμε πολλούς διαφορετικούς συνδυασμούς αριθμών νευρώνων για δίκτυα με δυο κρυφά επίπεδα. Για κάθε αρχιτεκτονική εκπαίδευσαμε 5000 ΤΝΔ
2. Φιλτράραμε τα εκπαιδευμένα αυτά δίκτυα εντοπίζοντας και αποθηκεύοντας εκείνο το οποίο είχε το μικρότερο σφάλμα ελέγχου
3. Για την τελική επιλογή της βέλτιστης αρχιτεκτονικής εντοπίσαμε αρχικά τις δομές με τα λιγότερα δεδομένα που είχαν σφάλμα >10% της ζητούμενης επιφανειακής τραχύτητας και στη συνέχεια προχωρήσαμε σε περαιτέρω διερεύνηση βασισμένη στη σύγκριση των μέσων τετραγωνικών σφαλμάτων

Όλα τα παραπάνω επαναλήφθηκαν για τρεις διαφορετικές μεθόδους-αλγόριθμους εκπαίδευσης:

- Τον προσαρμοστικό αλγόριθμο ανάστροφης διάδοσης της πιο απότομης κατάβασης με χρήση όρου ορμής "**trainGDX**"
- Τον αλγόριθμο ανάστροφης διάδοσης των Levenberg-Marquardt "**trainLM**"
- Τον αλγόριθμο συζυγών κλίσεων scaled conjugate gradient "**trainSCG**"

Για την καλύτερη σύγκλιση κατά την εκπαίδευση των ΤΝΔ, στα παραδείγματα εκπαίδευσης και ελέγχου του δικτύου εφαρμόστηκε η τεχνική της κανονικοποίησης δεδομένων σε ένα εύρος μεταξύ 0–1 διαιρώντας, για κάθε παράμετρο, όλες τις τιμές με την μέγιστη τιμή της παραμέτρου.

Τα βέλτιστα δίκτυα στα οποία καταλήξαμε για κάθε μια από τις τρεις μεθόδους εκπαίδευσης φαίνονται στον παρακάτω πίνακα:

Μέθοδος Εκπαίδευσης	Δομή	MSE Ελέγχου	MSE Εκπαίδευσης	Δεδομένα με σφάλμα >10%
trainGDX	3-19-10-1	2,29 E-04	4,65 E-04	1
trainLM	3-11-5-1	1,20 E-04	2,60 E-04	0
trainSCG	3-18-9-1	2,83 E-04	3,98 E-04	0

**Πίνακας 4.21** Βέλτιστα ΤΝΔ για κάθε μέθοδο εκπαίδευσης.

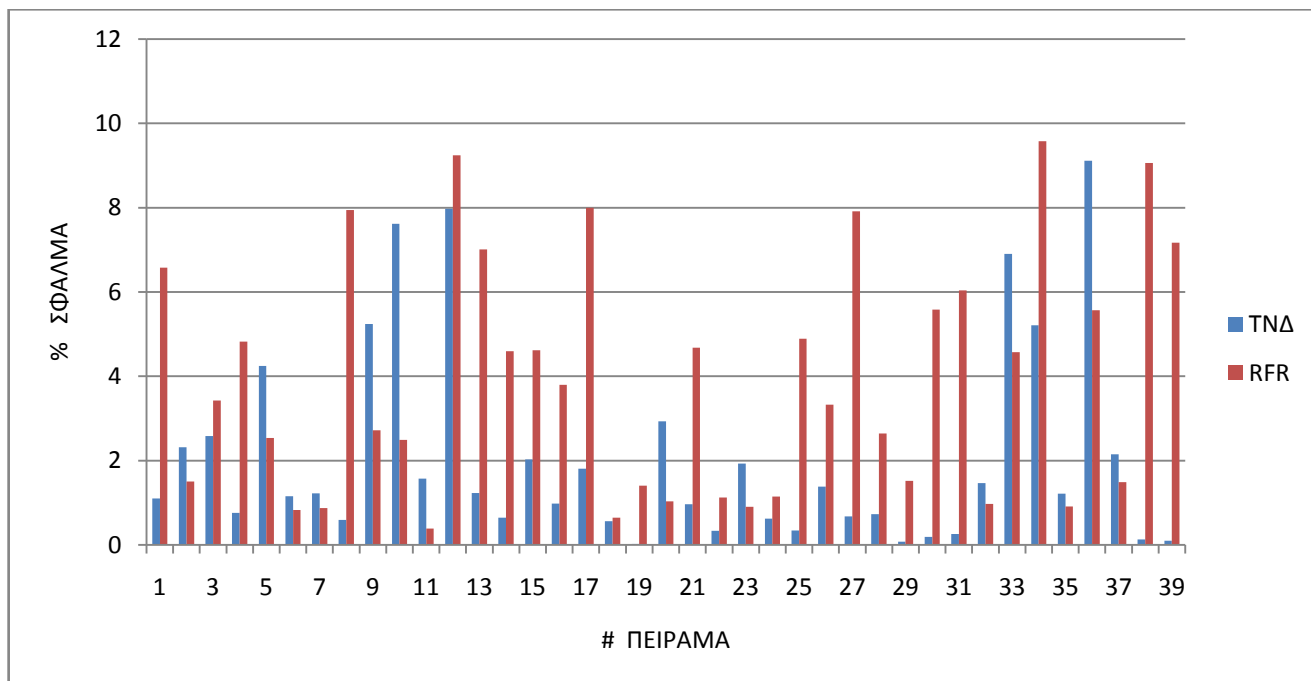
#### Παρατηρήσεις:

- Όσον αφορά το μέσο τετραγωνικό σφάλμα εκπαίδευσης, η μέθοδος trainGDX έδωσε τα χειρότερα αποτελέσματα από τις τρεις μεθόδους. Ανάμεσα στις μεθόδους trainLM και trainSCG, η μέθοδος trainLM έδωσε οριακά μικρότερα σφάλματα.
- Όσον αφορά το μέσο τετραγωνικό σφάλμα ελέγχου, η μέθοδος trainGDX έδωσε οριακά χειρότερα αποτελέσματα από τη μέθοδο trainLM, ενώ μεταξύ των μεθόδων trainLM και trainSCG διαπιστώνουμε και ότι η μέθοδος trainSCG έδωσε μικρότερα σφάλματα ελέγχου.
- Όσον αφορά τον ρόλο που παίζει το μέγεθος του δικτύου (ο αριθμός των νευρώνων στα κρυφά επίπεδα) στην απόδοσή του, παρατηρήσαμε ότι καθώς κινούμαστε προς την αύξηση του αριθμού νευρώνων στο πρώτο κρυφό επίπεδο, φαίνεται να αυξάνεται ελαφρώς το σφάλμα ελέγχου για όλες τις μεθόδους, ενώ για το σφάλμα εκπαίδευσης δεν παρατηρήθηκε κάποιο αντίστοιχο συμπέρασμα.

Από τα τρία βέλτιστα ΤΝΔ για κάθε μέθοδο εκπαίδευσης επιλέξαμε για σύγκριση με το μοντέλο παλινδρόμησης με τυχαία δάση, εκείνο το οποίο είχε το μικρότερο μέσο τετραγωνικό σφάλμα εκπαίδευσης και ελέγχου. Το δίκτυο αυτό είναι το βέλτιστο ΤΝΔ της μεθόδου trainLM με δομή 3-11-5-1. Το ΤΝΔ που

επιλέξαμε ως βέλτιστο δίνει σαφώς καλύτερα αποτελέσματα από το μοντέλο παλινδρόμησης με τυχαία δάση το οποίο, από τα συνολικά 39 πειράματα, δίνει οριακά καλύτερο αποτέλεσμα μόνο σε 10 πειράματα ενώ παρουσιάζει σημαντικά μεγαλύτερο σφάλμα σε σχέση με τα αποτελέσματα του TNΔ σε 8 δεδομένα.

Παρακάτω παρουσιάζεται ένα συγκριτικό διάγραμμα με το % σφάλμα του TNΔ της μεθόδου trainLM με δομή 3-11-5-1 και του μοντέλου τυχαίων δασών:



**Διάγραμμα 4.22** Συγκριτικό διάγραμμα % σφάλματος μεταξύ του TNΔ της μεθόδου trainLM με δομή 3-11-5-1 και του μοντέλου με τυχαία δάση.

Όπως φαίνεται και στο παραπάνω διάγραμμα κανένα από τα δύο μοντέλα δεν δίνει τιμές επιφανειακής τραχύτητας με σφάλμα που να ξεπερνά το 10% των πειραματικών μετρήσεων, αλλά το TNΔ παρέχει γενικά μικρότερο σφάλμα.

## ΠΑΡΑΡΤΗΜΑ

Παρακάτω παρατίθεται ο κώδικας που αναπτύχθηκε στη MATLAB για την κατασκευή, εκπαίδευση και αποθήκευση των ΤΝΔ:

```
%***** MULTI LAYER PERCEPTRON NEURAL NETWORK *****%
% Το παρακάτω πρόγραμμα δημιουργεί νευρωνικά δίκτυα 2 κρυφών επιπέδων i και j
% Για κάθε νευρωνικό [i j] θα γίνονται πολλά retrains και θα αποθηκεύεται τελικά το καλύτερο
retrain

% Εισαγωγή των πινάκων Feed, Depth, Speed, Roughness με τιμές από συγκεκριμένο paper
feed=[0.08;0.08;0.08;0.08;0.08;0.08;0.08;0.09;0.09;0.09;0.09;0.09;0.09;0.09;0.09;0.1;
0.1;0.1;0.1;0.1;0.1;0.1;0.1;0.1;0.1;0.1;0.15;0.15;0.15;
0.15;0.15;0.15;0.15;0.15;0.15;0.15;0.15;];
feed=feed./max(feed);
depth=[0.1;0.105;0.2;0.2;0.452;0.542;0.935;0.083;0.125;0.144;0.2;0.2;0.2;0.542;0.542;0.753;
0.935;0.045;0.048;0.133;0.2;0.2;0.234;0.352;0.542;0.558;0.754;
0.935;0.019;0.06;0.1;0.2;0.2;0.2;0.278;0.542;0.657;0.906;0.935];
depth=depth./max(depth);
speed=[1608;1250;858;965;1850;1072;1072;2145;1000;1072;858;965;1072;965;1072;2050;1072;2145;
2681;1608;858;965;2145;2220;1072;1400;858;1072;2681;1287;
2681;858;965;1072;1608;1072;1600;2600;1072];
speed=speed./max(speed);
roughness=[0.502;0.532;0.5902;0.539;0.592;0.5693;0.5821;0.667;0.735;0.683;0.6776;0.6179;0.742;
0.718;0.65;0.764;0.625;0.77;0.781;0.773;0.6687;0.7029;0.772;
0.784;0.6769;0.812;0.809;0.6966;1.251;
1.361;1.193;1.134;1.0854;1.316;1.312;1.1083;1.345;1.523;1.1337];
roughness=roughness./max(roughness);

% Προσδιορισμός των πινάκων Inputs και Targets, επιλογή συνδυασμών αριθμού νευρώνων σε κάθε ένα
από τα δυο hidden layers και επιλογή αριθμού retrains
% για κάθε νευρωνικό
inputs=[feed,depth,speed];
targets=roughness;
inputs=inputs';
targets=targets';
a1=5; %Αριθμός νευρώνων στο 1ο layer (από a1 έως b1)
b1=20;
a2=5; %Αριθμός νευρώνων στο 2ο Layer (από a2 έως b2)
b2=20;
c=5000; %Αριθμός retrains κάθε νευρωνικού δικτύου

for i=a1:1:b1 % Με αυτά τα 2 for δίνουμε ταυτότητα στο νευρωνικό δίκτυο με βάση τον
αριθμό νευρώνων του κάθε layer
for j=a2:1:b2

mseErrorNet(i,j)=1;

for z=1:1:c % Με αυτό το for τρέχουμε πολλά retrains για κάθε συγκεκριμένο νευρωνικό
netss(z).net = newff(inputs,targets,[i j],
('tansig','tansig','purelin'),'trainlm');
netss(z).net = init(netss(z).net);
netss(z).net.divideFcn='divideind';
netss(z).net.divideParam.trainInd = [1 2 3 6 7 8 9 12 13 14 17 18 19 20 21 22 25
26 27 28 29 32 33 34 35 38 39]; % 27 τιμές
netss(z).net.divideParam.valInd = [4 10 15 23 30 36]; % 6 τιμές
netss(z).net.divideParam.testInd = [5 11 16 24 31 37]; % 6 τιμές
[netss(z).net,tr(z)]=train(netss(z).net,inputs,targets);
outputs(z,:)=sim(netss(z).net,inputs);
mseError(z)=mean((targets(tr(z).testInd)-outputs(z,tr(z).testInd)).^2);
```

```

        if mseError(z)<mseErrorNet(i,j) % Αποθηκεύω από c διαφορετικά Trains το καλύτερο
με βάση το mse error των Test indices
            nets(i,j).net = netss(z).net;
            mseErrorNet(i,j)=mseError(z);
            outputsNet(i,j,:)=outputs(z,:);
            trains(i,j)=tr(z); % Αποθηκεύω τα δεδομένα του καλύτερου train για το
συγκεκριμένο νευρωνικό δίκτυο
            disp('beltisto neurwniko diktyo'),disp(z),disp(mseErrorNet(i,j)); % Εμφανίζω στο
command line τον αριθμό τον retrain και το mseError κάθε νέου βέλτιστου νευρωνικών

        else
            %disp('oxi beltisto'),disp(z); %na emfanizei sto command line an den proekipse
kalutero net
        end
    end
    MeanMseErrorNetRetrains(i,j) = sum(mseError)/c;
end
end

% Αποθηκεύω τα plot_OutputsTargets, plot_regression και plot_performance με την βοήθεια των
εντολών saveas που εμπεριέχονται στα αντίστοιχα for.
% Στα saveas αυτά υπάρχει μια διαδρομή της μορφής π.χ.: C:/Users/Manolis/Pictures/Outputs-
Targets/%dx%d.png
% Αυτό σημαίνει ότι πρέπει για να αποθηκευθεί το αντίστοιχο plot να υπάρχει ήδη ένας φάκελος με
την αντίστοιχη ονομασία (εδώ π.χ.: "Outputs-Targets") και να βρίσκεται
% στη διαδρομή: C:/Users/Manolis/Pictures/
% Έτσι θα αποθηκεύσει μια εικόνα με ονομασία: %dx%d.png όπου ο όρος %dx%d θα αντιστοιχεί στο
14x3 ή όποιον άλλον αριθμό νευρώνων τρέξουμε

% Για τη δημιουργία των διαγραμμάτων Regression ομαδοποιούμε τις μεταβλητές που περιέχουν τις
τιμές Train,
Validation και Test για το κάθε καλύτερο νευρωνικό των retrain ως "Targets" και "Outputs"
for i=a1:1:b1
    for j=a2:1:b2
        tsTarg(i,j,:) = targets(trains(i,j).testInd);
        tsOut(i,j,:) = reshape(outputsNet(i,j,(trains(i,j).testInd)),1,6);
        trTarg(i,j,:) = targets(trains(i,j).trainInd);
        trOut(i,j,:) = reshape(outputsNet(i,j,(trains(i,j).trainInd)),1,27);
        vTarg(i,j,:) = targets(trains(i,j).valInd);
        vOut(i,j,:) = reshape(outputsNet(i,j,(trains(i,j).valInd)),1,6);
    end
end

% Αποθηκεύω σε συγκεκριμένο φάκελο τα plot outputs+targets με τα παραδείγματα των βέλτιστων
νευρωνικών ως εικόνες .png με ονομασία των αριθμό των
% νευρώνων των hidden layers (π.χ. 14x3)
paradeigmata=1:39;
for i=a1:1:b1
    for j=a2:1:b2
        OutuptsTargetsFigures=figure;
        hold on
        % ylim([0.3 1.1])
        scatter(paradeigmata(trains(i,j).testInd), reshape(outputsNet(i,j,(trains(i,j).testInd)),
1,6),40,'r','+') % Τα test να έχουν κόκκινο χρώμα
        scatter(paradeigmata(trains(i,j).valInd), reshape(outputsNet(i,j,(trains(i,j).valInd)),
1,6),40,'g','+') % Τα validation να έχουν πράσινο χρώμα
    end
end

```



```

scatter(paradeigmata(trains(i,j).trainInd), reshape(outputsNet(i,j,(trains(i,j).trainInd)),
1,27),40,'b','*') % Τα train να έχουν μπλε χρώμα
scatter(paradeigmata(targets,'b','o')
saveas(OutuptsTargetsFigures,sprintf('C:/Users/Manolis/Pictures/Outputs-Targets/%dx%d.
png',i,j));
end
end

% Αποθηκεύω σε συγκεκριμένο φάκελο τα plotregression των βέλτιστων νευρωνικών ως εικόνες .png
με ονομασία των αριθμό των νευρώνων των hidden layers (π.χ. 14x3)
for i=a1:1:b1
for j=a2:1:b2
RegressionFigures=figure;
plotregression(reshape(trTarg(i,j,:),1,27), reshape(trOut(i,j,:),1,27), 'Train', reshape
(vTarg(i,j,:),1,6), reshape(vOut(i,j,:),1,6),
'Validation', reshape(tsTarg(i,j,:),1,6), reshape(tsOut(i,j,:),1,6), 'Testing');
saveas(RegressionFigures,sprintf('C:/Users/Manolis/Pictures/Regression/%dx%d.png',i,j));
%Prepei na einai katw apo fig kai na exei "/" anti gia "\" gia na leitourgisei
end
end

% Αποθηκεύω σε συγκεκριμένο φάκελο τα plot performance των βέλτιστων νευρωνικών ως εικόνες .png
με ονομασία των αριθμό των νευρώνων των hidden layers (π.χ. 14x3)
for i=a1:1:b1
for j=a2:1:b2
PerformanceFigures=figure;
plotperform(trains(i,j));
saveas(PerformanceFigures,sprintf('C:/Users/Manolis/Pictures/Performance/%dx%d.png',i,j));
end
end

```

## **ΒΙΒΛΙΟΓΡΑΦΙΑ**

- [1] A.Agrawal, S.Goel, W. Bin Rashid, M.Price (2014), Prediction of surface roughness during hard turning of AISI 4340 steel (69 HRC), Applied Soft Computing, 30, 279-286
- [2] Δ. Ε. Μανωλάκος (2011) Κατεργασίες II, Σημειώσεις μαθήματος, Ε.Μ.Π.
- [3] Μανταλία Γεωργία (2009), Διπλωματική εργασία με θέμα: Μελέτη των δυνάμεων κοπής στις κατεργασίες υψηλών ταχυτήτων, Αθήνα
- [4] Markus M.W. Knufermann (2003/2004), PhD thesis: Machining surfaces of optimal quality by hard turning , cranfield university, UK
- [5] Simon Haykin (1998), Neural Networks: A comprehensive foundation, 2nd edition, Prentice Hall, USA
- [6] Ι.Κ. Χατζηλάου, Δ. Κουπατσιάρης Π.Ν., Γ.Ι.Τσεκούρας (2008), Εισαγωγή στα νευρωνικά δίκτυα - Εργαστήρια Ηλεκτροτεχνίας, Σχολή Ναυτικών Δοκίμων
- [7] Ζαφειριάδου Ανδρονίκη - Κυνηγαλάκης Μύρων (2005), Διπλωματική εργασία με θέμα: Πρόβλεψη επιφανειακής τραχύτητας σε κατεργασία φρεζαρίσματος με κονδύλι με τη χρήση νευρωνικών, ασαφών και νευρο-ασαφών δικτύων, Αθήνα
- [8] Κ.Διαμαντάρας (2007), Τεχνητά Νευρωνικά Δίκτυα - Εκδόσεις Κλειδάριθμος, Αθήνα
- [9] B. Anuja Beatricea, E. Kirubakaranb, P. Ranjit Jeba Thangaiahc, K. Leo Dev Winsd (2014) , Surface Roughness Prediction using Artificial Neural Network in Hard Turning of AISI H13 Steel with minimal Cutting Fluid Application, Karunya University, India
- [10] J.Gerald Anto Arulraj, K.Leo Rec Wins, Anil Raj (2014), Artificial neural network assisted sensor fusion model for predicting surface roughness during hard turning of H13 steel with minimal cutting fluid application, , School of mechanical sciences, Karunya University, India
- [11] Hamdi Aouici a, Mohamed Athmane Yaltese, Kamel Chaoui, Tarek Mabrouki, Jean-Francois Rigal (2011), Analysis of surface roughness and cutting force components in hard turning with CBN tool: Prediction model and cutting conditions optimization, Algeria - France
- [12] Tug̃rul Ozel, Yig̃it Karpat (2004) Predictive modeling of surface roughness and tool wear in hard turning using regression and neural networks, Department of Industrial and Systems Engineering, Rutgers, The State University of New Jersey, USA
- [13] İlhan Asiltörk, Mehmet Hunkas (2010), Modeling and prediction of surface roughness in turning operations using artificial neural network and multiple regression method, Turkey
- [14] Mark Hudson Beale, Martin T. Hagan , Howard B. Demuth (2015), MATLAB Neural Network Toolbox (Users Guide) R2015a ,USA
- [15] Mark Hudson Beale, Martin T. Hagan , Howard B. Demuth (2015), MATLAB Neural Network Toolbox (Reference) R2015a, USA
- [16] <http://www.mathworks.com/help/nnet/ug/improve-neural-network-generalization-and-avoid-overfitting.html#bt0cnqi>

- [17] Ν.Γαλάνης (2011) Διδακτορική διατριβή: Κατεργασιμότητα λίαν υψηλής ακρίβειας βιολογικών εμφυτευμάτων τεχνητού ισχίου, Ε.Μ.Π., Αθήνα
- [18] Σ.Τζαφέστας (2002), Υπολογιστική Νοημοσύνη, Τόμος Α: Μεθοδολογίες, Αθήνα
- [19] Σ.Τζαφέστας (2002), Υπολογιστική Νοημοσύνη, Τόμος Β: Εφαρμογές, Αθήνα
- [20] Γ.Ρίζος (1996), Τεχνητά Νευρωνικά δίκτυα - Εκδόσεις νέων τεχνολογιών, Αθήνα
- [21] Λ.Ε.Λαζαρίδη (1997), Μηχανουργική τεχνολογία - Ίδρυμα Ευγενίδου, Αθήνα
- [22] Ι. Ρεφανίδης (2011), Ηλεκτρονικές Διαφάνειες "Νευρωνικά Δίκτυα"
- [23] <http://www.mathworks.com/products/neural-network/code-examples.html>
- [24] Sung Hoon Oh (2014), A Study on Cutting Force Characteristics in Hard Turning, Division of Mechanical system engineering, Chon-buk National University
- [25] Suha Karim Shihab, Zahid A. Khan, Aas Mohammad & Arshad Noor Siddiquee (2014), A review of turning of hard steels used in bearing and automotive applications, Production & Manufacturing Research, 2:1, 24-49, India
- [26] V.P. Bhemuni and S.R. Chalamalasetti (2013), A Review on Hard Turning by using Design of Experiments, India
- [27] E. García-Plazaa, P.J. Núñez, D.R. Salgadob, I. Camberob, J.M. Herrera Olivenzab, J. García Sanz-Calcedob (2013), Surface Finish Monitoring in Taper Turning CNC Using Artificial Neural Network and Multiple Regression Methods, Technical School of Industrial Engineering- Department of Applied Mechanics and Engineering Projects, University of Castilla-La Mancha, Spain
- [28] Ravinder Kumar, Santram Chauhan (2013), Study on surface roughness measurement for turning of Al 7075/10/SiCp and Al 7075 hybrid composites by using response surface methodology (RSM) and artificial neural networking (ANN), Department of Mechanical Engineering, NIT Delhi