



**ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ**

**ΣΧΟΛΗ ΜΗΧΑΝΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ**

**ΔΙΕΡΕΥΝΗΣΗ ΤΗΣ ΑΠΟΔΟΣΗΣ ΜΟΝΤΕΛΩΝ  
ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ ΚΑΤΑ ΤΗΝ  
ΠΡΟΒΛΕΨΗ ΤΗΣ ΕΠΙΦΑΝΕΙΑΚΗΣ  
ΤΡΑΧΥΤΗΤΑΣ ΣΤΗΝ ΚΑΤΕΡΓΑΣΙΑ ΤΗΣ  
ΛΕΙΑΝΣΗΣ**

**ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ ΤΟΥ**

**ΠΑΠΑΔΟΠΟΥΛΟΥ ΙΩΑΝΝΗ**

**ΕΠΙΒΛΕΠΩΝ**

**ΛΕΚΤΟΡΑΣ Α. ΜΑΡΚΟΠΟΥΛΟΣ**

Αθήνα, Μάρτιος 2016





**NATIONAL TECHNICAL UNIVERSITY OF ATHENS**  
**SCHOOL OF MECHANICAL ENGINEERING**

**INVESTIGATION OF THE PERFORMANCE OF  
ARTIFICIAL NEURAL NETWORK MODELS FOR  
THE PREDICTION OF SURFACE ROUGHNESS  
DURING GRINDING**

**DIPLOMA THESIS OF**

PAPADOPOULOS IOANNIS

**SUPERVISOR**

LECTURER A. MARKOPOULOS

Athens, March 2016



**Θερμές ευχαριστίες εκφράζονται στον Λέκτορα κ. Α.  
Μαρκόπουλο και στον Υποψήφο Διδάκτορα κ. Ν. Κάρκαλο  
για την καθοδήγηση κατά την διάρκεια συγγραφής της  
παρούσας εργασίας**



## Περιεχόμενα

Περίληψη .....	1
1 Κατεργασία Λείανσης μετάλλων .....	2
1.1 Εισαγωγή .....	2
1.2 Είδη λείανσης.....	2
1.2.1 Επιφανειακή Λείανση (surface grinding) .....	2
1.2.2 Κυλινδρική Λείανση (Cylindrical Grinding).....	3
1.2.3 Έκκεντρη λείανση (Centerless Grinding).....	6
1.2.4 Λείανση μεγάλης Πρόωσης (Creep Feed Grinding).....	7
1.3 Συνθήκες λείανσης.....	8
1.3.1 Τελείωμα επιφάνειας .....	9
1.3.2 Επίδραση θερμοκρασίας .....	9
1.3.3 Φθορά λειαντικού τροχού .....	10
1.4 Λειαντικός τροχός.....	10
1.4.1 Υλικό κόκκων λείανσης.....	11
1.4.2 Μέγεθος κόκκων λείανσης .....	12
1.4.3 Συνδετικό υλικό κόκκων λειαντικού τροχού .....	12
1.4.4 Σκληρότητα (grade) λειαντικού τροχού.....	14
1.4.5 Υφή (structure) λειαντικού τροχού .....	14
1.4.6 Γεωμετρία τροχού λείανσης .....	15
1.4.7 Προδιαγραφές και κωδικοποίηση τροχών λείανσης.....	15
1.5 Μηχανική Λείανσης.....	16
1.5.1 Επίπεδη περιφερειακή λείανση.....	16
1.5.2 Κυλινδρική λείανση .....	19
1.5.2.1 Μέγιστο βάθος κοπής ανά κόκκο .....	19
1.5.3 Δυνάμεις και ισχύς λείανσης .....	21
1.5.4 Δρώσα δύναμη ανά κόκκο .....	22
2. Νευρωνικά δίκτυα.....	23
2.1 Εισαγωγή .....	23
2.2 Λειτουργία των νευρωνικών δικτύων .....	24
2.3 Αρχιτεκτονικές των νευρωνικών δικτύων .....	26
2.3.1 Δίκτυα πρόσθιας τροφοδότησης ενός επιπέδου.....	27
2.3.2 Δίκτυα πρόσθιας τροφοδότησης πολλαπλών επιπέδων.....	27
2.3.3 Δίκτυα με ανάδραση .....	29
2.4 Μορφές της συνάρτησης ενεργοποίησης των νευρωνικών δικτύων .....	30

2.4.1 Βηματική συνάρτηση (Heaviside function).....	30
2.4.2 Γραμμική συνάρτηση.....	31
2.4.3 Γραμμική συνάρτηση ξεχωριστών βημάτων .....	31
2.4.4 Σιγμοειδής συνάρτηση .....	32
2.5 Εκπαίδευση Νευρωνικών Δικτύων .....	34
2.5.1 Νόμοι εκπαίδευσης .....	35
2.5.2 Αλγόριθμοι εκπαίδευσης .....	36
2.6 Ικανότητα γενίκευσης αποτελεσμάτων νευρωνικών δικτύων .....	43
3. Νευρωνικά δίκτυα ακτινικής βάσης (radial basis function neural networks).....	45
3.1 Δομή και τρόπος λειτουργίας των νευρωνικών δικτύων ακτινικής βάσης...	46
3.2 Εκπαίδευση νευρωνικών δικτύων ακτινικής βάσης .....	53
4. Μαθηματική μοντελοποίηση λείανσης μετάλλων με χρήση νευρωνικών δικτύων	57
4.1 Πειραματικά δεδομένα.....	57
4.2 Κατασκευή νευρωνικών δικτύων πρόσθιας τροφοδότησης .....	64
4.3 Κατασκευή νευρωνικών δικτύων ακτινικής βάσης .....	66
5. Αποτελέσματα.....	68
5.1 Αποτελέσματα νευρωνικών δικτύων πρόσθιας τροφοδότησης.....	68
5.1.1 Επίδραση Αλγορίθμου εκπαίδευσης.....	74
5.1.2 Επίδραση Αρχιτεκτονικής Νευρωνικού δικτύου .....	78
5.2 Αποτελέσματα νευρωνικών δικτύων ακτινικής βάσης.....	83
5.2.1 Επίδραση Αρχιτεκτονικής νευρωνικών δικτύων ακτινικής βάσης.....	84
6. Συμπεράσματα/Προτάσεις για μελλοντική έρευνα .....	86
Βιβλιογραφία .....	88



## Περίληψη

Στην παρούσα εργασία έγινε απόπειρα μαθηματικής μοντελοποίησης της κατεργασίας λείανσης μετάλλων με χρήση νευρωνικών δικτύων πρόσθιας τροφοδότησης, καθώς επίσης και νευρωνικών δικτύων ακτινικής βάσης. Τα πειραματικά δεδομένα ελήφθησαν από την βιβλιογραφία και αφορούν την λείανση τριών ειδών χάλυβα με έξι διαφορετικούς τύπους λειαντικών τροχών. Για την προσέγγιση του προβλήματος, δοκιμάστηκε ένας μεγάλος αριθμός διαφορετικών αρχιτεκτονικών νευρωνικών δικτύων πρόσθιας τροφοδότησης, όπως επίσης και τεσσάρων διαφορετικών αλγορίθμων εκπαίδευσης. Επίσης δοκιμάστηκε ένας μεγάλος αριθμός αρχιτεκτονικών νευρωνικών δικτύων ακτινικής βάσης. Από τις δοκιμές που έγιναν, βρέθηκε ότι για την στατιστική μοντελοποίηση της συγκεκριμένης κατεργασίας, είναι προτιμότερη η χρήση των νευρωνικών δικτύων ακτινικής βάσης, καθώς το σφάλμα πρόβλεψης βρέθηκε να είναι μικρότερο. Εκτός αυτού, τα νευρωνικά δίκτυα ακτινικής βάσης έχουν χαμηλότερο υπολογιστικό κόστος, και απλούστερη αρχιτεκτονική, καθιστώντας τα πιο εύχρηστα στην προσέγγιση του συγκεκριμένου προβλήματος.

## ***Abstract***

*In the current thesis, mathematical modeling for the prediction of surface roughness during metal grinding was attempted, using feed forward and radial basis neural networks. Experimental data were obtained from relevant literature, which correspond to grinding of three types of steels, using six different types of grinding wheels. A large number of different feed forward and radial basis neural networks architectures were tested, and for the case of the feed forward neural networks, four different training algorithms were compared. Computational results have shown that mathematical models based on radial basis neural networks, were able to predict the experimental data with smaller errors, compared to feed forward neural networks predictions. Moreover, due to their simpler architecture and smaller computational cost, they proved advantageous during the approach of the specific problem.*

# 1 Κατεργασία Λείανσης μετάλλων

## 1.1 Εισαγωγή

Με τον όρο λείανση εννοείται η κατεργασία αποπεράτωσης ενός υλικού, η οποία περιλαμβάνει την αφαίρεση μέρους αυτού. Η κατεργασία της λείανσης εφαρμόζεται συνήθως όταν συντρέχουν οι παρακάτω λόγοι:

- Το υπό επεξεργασία υλικό παρουσιάζει μεγάλη σκληρότητα
- Το υπό επεξεργασία υλικό παρουσιάζει αυξημένη ψαθυρότητα
- Η επιθυμητή επιφανειακή ποιότητα είναι υψηλή
- Η επιθυμητή ακρίβεια σε διαστάσεις είναι υψηλή.

Για την πραγματοποίηση της λείανσης χρησιμοποιούνται λειαντικοί κόκκοι σε τροχό λείανσης (bonded abrasive tools). Οι κόκκοι που χρησιμοποιούνται, έχουν ακαθόριστη γεωμετρία.

Σε όλες τις περιπτώσεις λείανσης, πραγματοποιείται αφαίρεση τμήματος του υλικού προς κατεργασία.

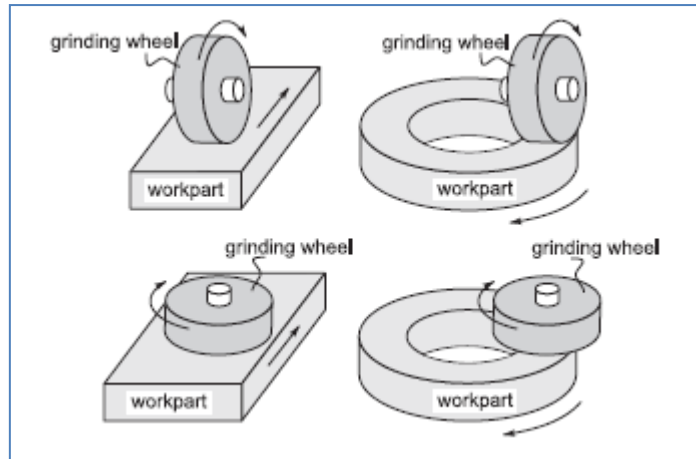
## 1.2 Είδη λείανσης

Η κατεργασία της λείανσης είναι δυνατόν να κατηγοριοποιηθεί σαν επιφανειακή, κυλινδρική, έκκεντρη και λείανση μεγάλης πρόωσης. Τα είδη αυτά της λείανσης αναλύονται στις παρακάτω παραγράφους.

### 1.2.1 Επιφανειακή λείανση (surface grinding)

Επιφανειακή λείανση είναι μία κατεργασία κατά την οποία ο λειαντικός τροχός κινείται στην επίπεδη επιφάνεια ενός υλικού και αφαιρεί ένα στρώμα του, όπως φαίνεται στο παρακάτω σχήμα.

Στην επιφανειακή λείανση, ο άξονας του λειαντικού τροχού είναι δυνατόν να είναι οριζόντιος ή κατακόρυφος, και η σχετική κίνηση μεταξύ τροχού και τεμαχίου να είναι παλινδρομική, ή περιστροφική, όπως φαίνεται στο παρακάτω σχήμα:



Σχήμα 1: Τέσσερις τύποι επιφανειακής λείανσης [1]

Από τους τέσσερις τύπους που παρουσιάζονται στο παραπάνω σχήμα, το συνηθέστερο είναι με λειαντικό τροχό οριζοντίου άξονα περιστροφής και παλινδρομική τράπεζα για το τεμάχιο προς λείανση (βλ. Σχήμα 2).

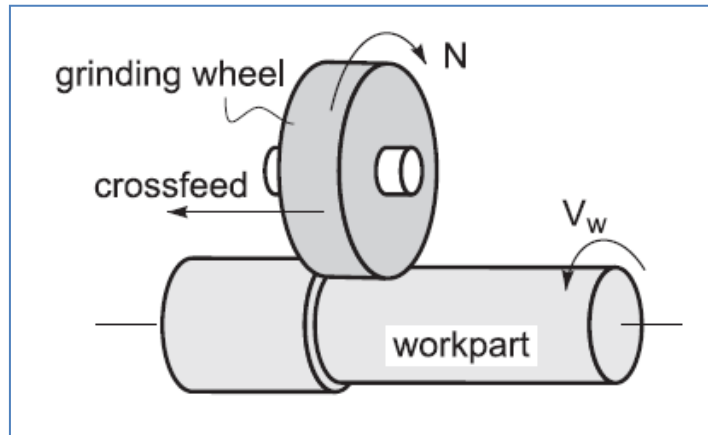


Σχήμα 2: Μηχανή λείανσης με λειαντικό τροχό οριζόντιου άξονα περιστροφής και παλινδρομική τράπεζα [1]

### 1.2.2 Κυλινδρική Λείανση (Cylindrical Grinding)

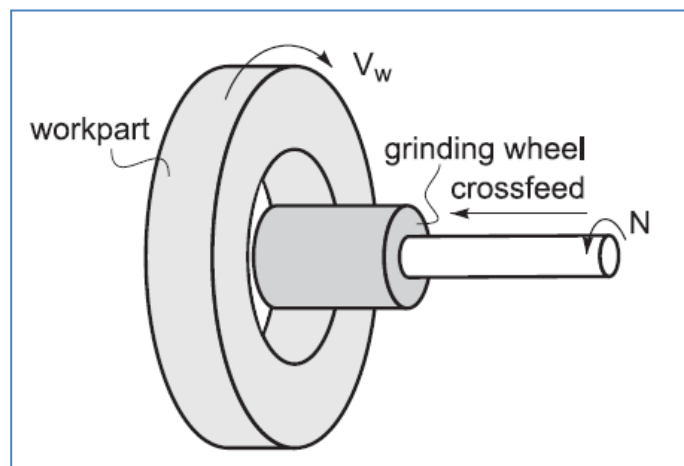
Κατά την κατεργασία της κυλινδρικής λείανσης, το τεμάχιο περιστρέφεται γύρω από τον άξονά του, και επίσης εκτελεί παλινδρομική κίνηση κατά μήκος του άξονά του.

Όταν ο λειαντικός τροχός λειαίνει την εξωτερική του επιφάνεια, το είδος αυτό της κατεργασίας ονομάζεται εξωτερική κυλινδρική λείανση, όπως φαίνεται στο παρακάτω σχήμα.



Σχήμα 3: Εξωτερική κυλινδρική λείανση [1]

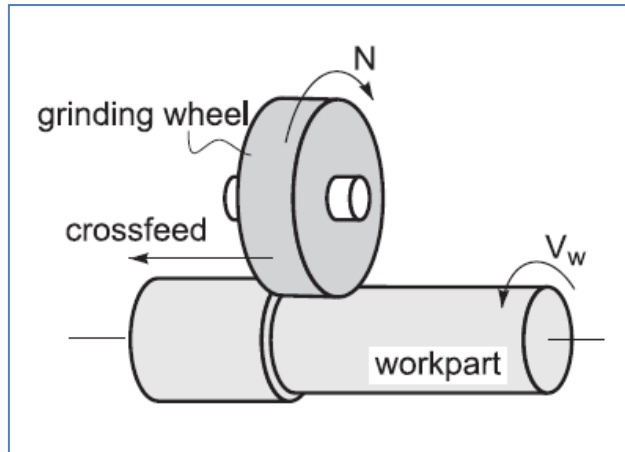
Όταν αντίθετα ο λειαντικός τροχός λειαίνει την εσωτερική επιφάνεια ενός τεμαχίου, αυτό ονομάζεται εσωτερική κυλινδρική λείανση (βλ. Σχήμα 4).



Σχήμα 4: Εσωτερική κυλινδρική λείανση [1]

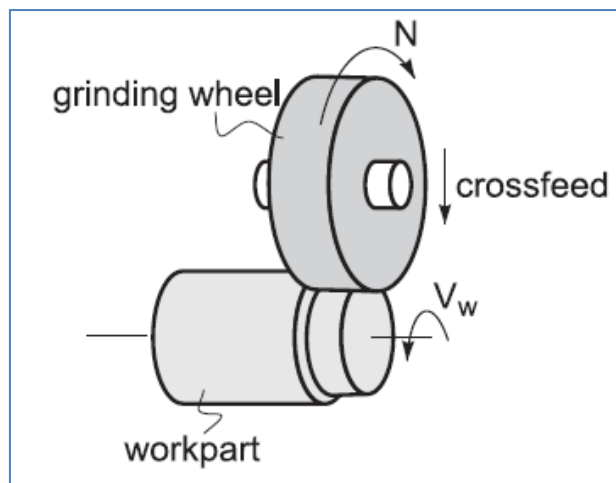
Για την κυλινδρική λείανση υπάρχουν τρεις τύποι πλευρικής πρόωσης του λειαντικού τροχού σε σχέση με το τεμάχιο:

- Παράλληλη πλευρική πρόωση με τον άξονα συμμετρίας του τεμαχίου (traverse feed grinding), βλ. Σχήμα.



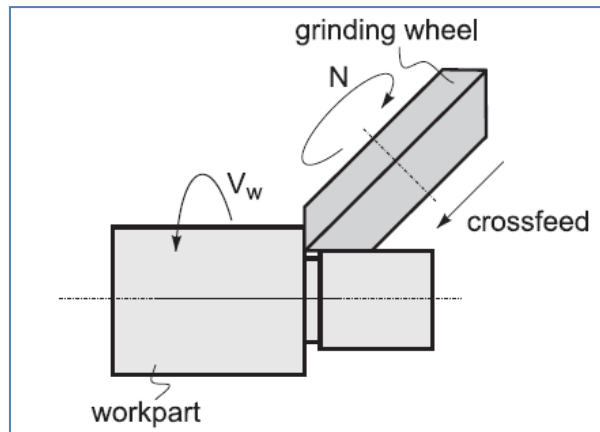
Σχήμα 5: Κυλινδρική Λείανση με παράλληλη πρόωση [1]

- Κάθετη πλευρική πρόωση σε σχέση με τον άξονα συμμετρίας του τεμαχίου (plunge grinding)



Σχήμα 6: Κυλινδρική λείανση με κάθετη πρόωση τεμαχίου [1]

- Συνδυασμός των δύο παραπάνω



Σχήμα 7: Συνδυασμός παράλληλης και κάθετης πρόωσης [1]

Τονίζεται εδώ ότι είναι εφικτό, μία μηχανή κυλινδρικής λείανσης να επιτρέψει όλα τα παραπάνω είδη κυλινδρικής λείανσης (universal cylindrical grinder).

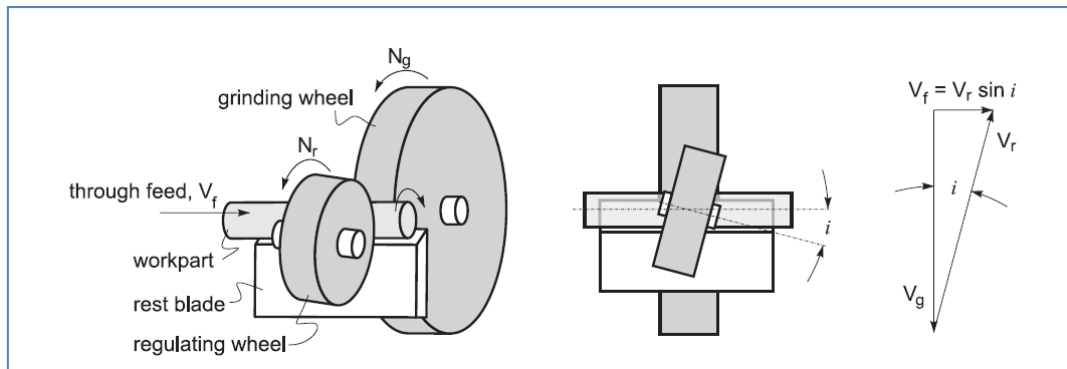


Σχήμα 8: Μηχανή κυλινδρικής λείανσης (Universal Cylindrical Grinder)

### 1.2.3 Έκκεντρη λείανση (Centerless Grinding)

Κατά την έκκεντρη λείανση, το τεμάχιο δεν στηρίζεται στον άξονά του ή σε κάποια οριζόντια επιφάνεια, αλλά σε μία λεπίδα. Το τεμάχιο ευρίσκεται μεταξύ δύο τροχών διαφορετικού μεγέθους. Ο μεγαλύτερος τροχός πραγματοποιεί την λείανση, ενώ ο μικρότερος κανονίζει την ταχύτητα του τεμαχίου  $V$ .

Η έκκεντρη λείανση, είναι δυνατόν, όπως και η κυλινδρική να είναι εσωτερική ή εξωτερική, με παράλληλη ή κάθετη πλευρική ταχύτητα του λειαντικού τροχού. Η πιο συνήθης εκδοχή της έκκεντρης λείανσης παρουσιάζεται στο παρακάτω σχήμα:



Σχήμα 9: Διάταξη έκκεντρης λείανσης [1]

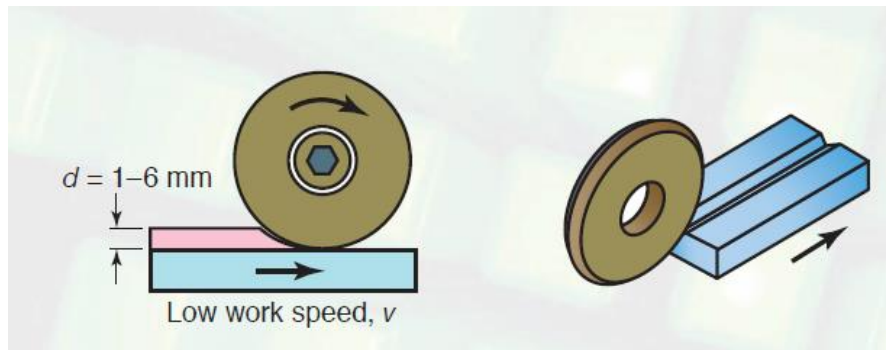
Οι μηχανές που πραγματοποιούν έκκεντρη λείανση είναι ιδανικές για μαζική παραγωγή και είναι δυνατόν να εκτελέσουν μόνο έναν τύπο έκκεντρης λείανσης.



Σχήμα 10: Μηχανή έκκεντρης λείανσης [1]

#### 1.2.4 Λείανση μεγάλης Πρόωσης (Creep Feed Grinding)

Κατά τα συνήθη είδη λείανσης, αφαιρείται μικρό στρώμα από το υλικό του τεμαχίου, ενώ υπάρχουν μεγάλες γωνιακές ταχύτητες στον τροχό λείανσης. Κατά την κατεργασία της λείανσης μεγάλης πρόωσης, το βάθος κοπής είναι δυνατόν να είναι αρκετά υψηλό, έως και 6mm με χαμηλή ταχύτητα τεμαχίου (βλ. σχήμα 11).

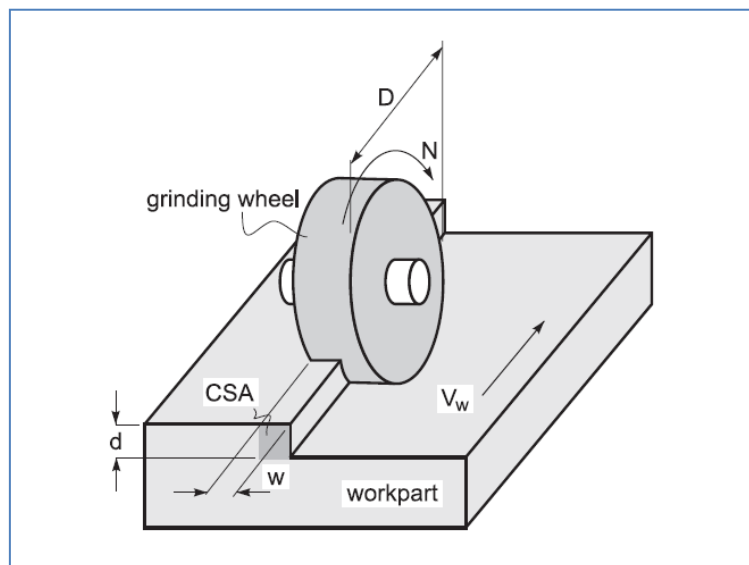


Σχήμα 11: Διάταξη λείανσης μεγάλης πρόωσης [2]

Ο τροχός λείανσης σε αυτήν την περίπτωση είναι μικρότερης σκληρότητας από τις άλλες κατηγορίες λείανσης.

### 1.3 Συνθήκες λείανσης

Η γεωμετρία της λείανσης απεικονίζεται στο παρακάτω σχήμα:



Σχήμα 12: Γεωμετρία λείανσης [1]

Η ταχύτητα κοπής  $v$  σε έναν τροχό λείανσης είναι εξαιρετικά υψηλή και συνδέεται με την περιστροφική ταχύτητα του τροχού, η οποία ορίζεται από την σχέση:

$$V = \pi DN$$



όπου:

$V$ : Η ταχύτητα κοπής

$D$ : Η διάμετρος του τροχού

$N$ : Η γωνιακή ταχύτητα του τροχού λείανσης

Βάθος κοπής ορίζεται ως η απόσταση μεταξύ της κατεργασμένης και της ακατέργαστης επιφάνειας  $d$  (βλ. Σχήμα 12).

Ταχύτητα κατεργαζόμενου τεμαχίου  $V_w$  είναι η περιφερειακή γραμμική ταχύτητα στην περίπτωση της κυλινδρικής λείανσης ενώ στην περίπτωση της επιφανειακής λείανσης είναι η ταχύτητα της παλινδρομικής κίνησης της τράπεζας.

### 1.3.1 Τελείωμα επιφάνειας

Το τελείωμα της επιφάνειας εξαρτάται από τις παρακάτω παραμέτρους:

- **Μέγεθος κόκκων λείανσης:** Μικρότεροι κόκκοι παράγουν επιφάνειες με μικρότερη τραχύτητα
- **Υφή:** Περισσότερο πυκνή διάταξη κόκκων λείανσης βελτιώνουν το τελείωμα του λειαινόμενου τεμαχίου, μειώνοντας την τραχύτητα
- **Ταχύτητα κοπής:** Με την αύξηση της ταχύτητας κοπής είναι δυνατόν να μειωθεί η τραχύτητα για το λειαινόμενο τεμάχιο.

Η ελάχιστη δυνατή τραχύτητα που είναι δυνατόν να επιτευχθεί μέσω λείανσης είναι της τάξεως των  $0.2\mu\text{m}$ .

### 1.3.2 Επίδραση θερμοκρασίας

Στην παράγραφο 1.3.1 αναφέρθηκαν παράγοντες οι οποίοι σχετίζονται με την γεωμετρία της κατεργασίας της λείανσης. Εκτός όμως από την γεωμετρία της λείανσης, σημαντικό ρόλο διαδραματίζει και η θερμοκρασία. Η αύξηση θερμοκρασίας λόγω τριβής είναι δυνατόν να επηρεάσει σημαντικά τις ιδιότητες της λειαινόμενης επιφάνειας, όπως επίσης και να προκαλέσει την θερμική διαστολή της, εισάγοντας διαστασιακά σφάλματα. Η αύξηση της θερμοκρασίας είναι δυνατόν να προκαλέσει τα παρακάτω [3]:

- Αυξημένες θερμοκρασίες είναι δυνατόν να προκαλέσουν μείωση στην σκληρότητα της λειαινόμενης επιφάνειας (tempering)

- Αν η θερμοκρασίες είναι εξαιρετικά υψηλές, είναι δυνατόν να προκληθεί καύση του υλικού. Η καύση προκαλεί ένα κυανό χρώμα σε χάλυβες, το οποίο δείχνει οξείδωση σε υψηλή θερμοκρασία
- Υψηλές θερμοκρασίες είναι δυνατόν να προκαλέσουν ρωγμές στο λειαινόμενο τεμάχιο. Οι ρωγμές είναι συνήθως κάθετες στην κατεύθυνση λείανσης, αλλά υπάρχουν και ακραίες περιπτώσεις όπου είναι δυνατόν να δημιουργηθούν παράλληλες ρωγμές.

Από τα παραπάνω, καθίσταται φανερό ότι η υπερβολική αύξηση θερμοκρασίας είναι δυνατόν να έχει αρνητική επίδραση στην επιφάνεια λείανσης. Κατά συνέπεια οι παράμετροι της κατεργασίας της λείανσης είναι απαραίτητο να επιλεγούν με ιδιαίτερη προσοχή. Επίσης η θερμοκρασία, είναι δυνατόν να ελεγχθεί με υγρά ψύξης.

### 1.3.3 Φθορά λειαντικού τροχού

Οι λειαντικοί τροχοί είναι δυνατόν να υποστούν τριών ειδών φθορές κατά την κατεργασία της λείανσης [3]:

- Θραύση κόκκων (grain fracture): Πραγματοποιείται όταν ένα μέρος του κόκκου αποσπάται, ενώ το εναπομείναν τμήμα του παραμένει προσκολλημένο στον τροχό. Οι γωνίες των κόκκων από τους οποίους αποσπάστηκαν θραύσματα είναι αιχμηρές, οπότε αποτελούν τις νέες επιφάνειες κοπής.
- Στόμωμα κόκκων (attritious wear): Περιλαμβάνει λείανση της επιφάνειας των κόκκων, το οποίο προκαλεί επίπεδες και εξομαλυμένες περιοχές.
- Διάσπαση συνδετικού υλικού κόκκων: Συμβαίνει όταν οι κόκκοι αποκολλώνται από το συνδετικό υλικό. Αυτό συμβαίνει συνήθως όταν οι κόκκοι έχουν λειανθεί πολύ (στόμωμα), και οι δυνάμεις κοπής είναι εξαιρετικά υψηλές. Όταν οι κόκκοι είναι σε καλή κατάσταση, οι δυνάμεις κοπής δεν είναι απαραίτητο να είναι υψηλές, για να επιτευχθεί το ίδιο αποτέλεσμα.

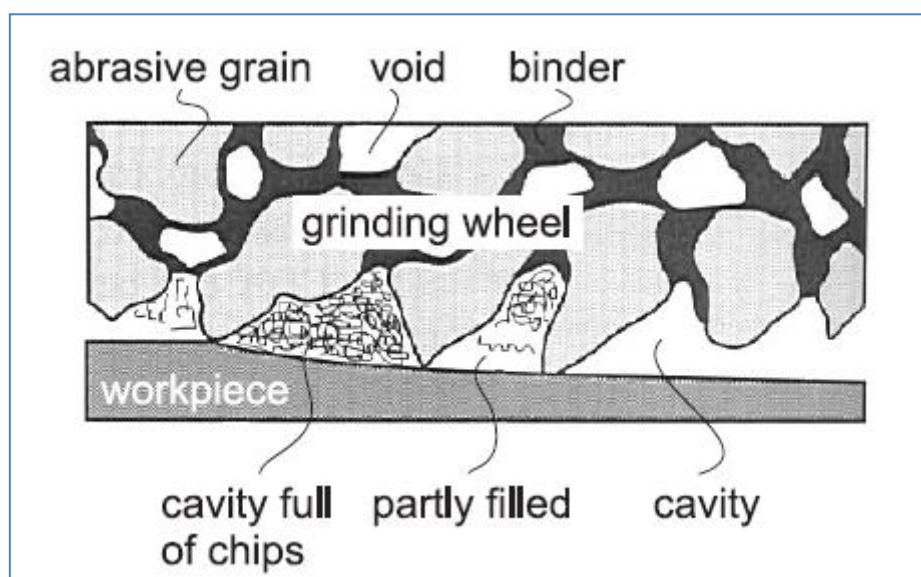
## 1.4 Λειαντικός τροχός

Ο λειαντικός τροχός αποτελείται από τους κόκκους, και το συνδετικό υλικό, χάρη στο οποίο, ο τροχός διατηρεί το σχήμα και την υφή του (βλ. Σχήμα 13).

Ο τρόπος με τον οποίο οι κόκκοι λείανσης, το συνδετικό υλικό, καθώς επίσης και ο κενός χώρος μεταξύ τους είναι δομημένοι, καθορίζουν τις παραμέτρους του λειαντικού τροχού, οι οποίες είναι:

- Υλικό κόκκων λείανσης
- Μέγεθος κόκκων λείανσης
- Συνδετικό υλικό κόκκων λείανσης
- Σκληρότητα (grade)
- Υφή (structure)
- Γεωμετρία τροχού λείανσης

Για να επιτευχθεί το επιθυμητό αποτέλεσμα κατά την κατεργασία της λείανσης, η επιλογή του σωστού τροχού είναι εξαιρετικά σημαντικός παράγοντας.



Σχήμα 13: Δομή τροχού λείανσης [1]

### 1.4.1 Υλικό κόκκων λείανσης

Τα συνηθέστερα συμβατικά υλικά για κόκκους λείανσης που χρησιμοποιούνται, είναι τα παρακάτω [3]:

- **Φυσικό ή τεχνητό κορούνδιο (Aluminium Oxide, Al<sub>2</sub>O<sub>3</sub>):** Πρόκειται για το πλέον σύνηθες μη μεταλλικό υλικό μεγάλης αντοχής
- **Ανθρακοπυρίτιο (Silicon Carbide, SiC):** Πρόκειται για σκληρότερο υλικό από το κορούνδιο, αλλά δεν είναι δυνατόν να χρησιμοποιηθεί σε μέταλλα

όπως χαλκός και αλουμίνιο, ανοξειδωτο χάλυβα, χυτοσίδηρο και αεροπορικά υλικά

Όταν οι απαιτήσεις είναι εξαιρετικά υψηλές, είναι δυνατή η ανάγκη για υπέρ-σκληρά υλικά, οπότε και χρησιμοποιούνται τα εξής;

- **Κυβικός βοριονιτρίτης (Cubic Boron Nitride – CBN):** Πρόκειται για εξαιρετικά σκληρό και ακριβό υλικό. Χρησιμοποιείται για κατεργασία σκληρυμένων μετάλλων εργαλείων, καθώς επίσης και σε αεροπορικά κράματα
- **Διαμάντι (Synthetic Diamond):** Ακόμα σκληρότερο και ακριβότερο από το παραπάνω. Χρησιμοποιείται για σκληρά και ψαθυρά υλικά όπως κεραμικά, καρβίδια και γυαλί.

#### 1.4.2 Μέγεθος κόκκων λείανσης

Το μέγεθος των κόκκων λείανσης είναι ένας σημαντικός παράγοντας για την τελική ποιότητα της επιφάνειας του τεμαχίου. Μικρά μεγέθη κόκκων δίνουν καλύτερο τελείωμα στην επιφάνεια, ενώ μεγάλα μεγέθη κόκκων αυξάνουν τον ρυθμό αφαίρεσης ύλης.

Οι κόκκοι λείανση που χρησιμοποιούνται στους λειαντικούς τροχούς είναι τυπικά μεταξύ 6 και 600, όπου 6 αντιστοιχεί στο μέγιστο δυνατό μέγεθος κόκκου και 600 στο ελάχιστο, όπως φαίνεται στον παρακάτω πίνακα [3]:

Χαρακτηρισμός τροχού	Κόκκωση τροχού				
Πολύ χονδρόκοκκος	8	10	12		
Χονδρόκοκκος	<u>14</u>	<u>16</u>	20	<u>24</u>	<u>30</u>
Μέσης κόκκωσης	36	40	<u>46</u>	50	<u>60</u>
Λεπτόκοκκος	70	<u>80</u>	90	100	120
Πολύ λεπτόκοκκος	150	<u>180</u>	200	220	<u>240</u>
Κόκκωση τύπου κόνεως	<u>260</u>	300	400	<u>500</u>	600

Πίνακας 1: Κωδικοποίηση μεγέθους κόκκων λείανσης [3]

#### 1.4.3 Συνδεδετικό υλικό κόκκων λειαντικού τροχού

Τα υλικά που είναι δυνατόν να χρησιμοποιηθούν σαν συνδεδετικά στους λειαντικούς τροχούς ανήκουν στις παρακάτω κατηγορίες [3]

- **Κεραμικά υλικά (Vitrified):** Είναι υαλώδη, ψαθυρά κα υψηλής αντοχής). Είναι κατάλληλα για ξηρή και υγρή λείανση και αντιπροσωπεύουν το 75% του συνόλου των χρησιμοποιούμενων λειαντικών τροχών.
- **Υδρύαλος (Silicate):** Είναι κατάλληλος για τρόχιση ειδικών εργαλείων, όπου η παραγωγή θερμότητας είναι απαραίτητο να ελαχιστοποιηθεί και η φθορά του τροχού δεν είναι σημαντικός παράγοντας
- **Συνθετικές Ρητίνες (Resinoid):** Χρησιμοποιείται σε λείανση μεγάλης ταχύτητας όπου προέχει ο υψηλός ρυθμός αφαίρεσης υλικού, όπως επίσης και στους τροχούς κοπής
- **Ρητινικό υλικό ενισχυμένο με ίνες (Resinoid Reinforced)**
- **Ελαστικά υλικά υψηλής αντοχής (Rubber):** Προσφέρονται για λείανση μεγάλης ταχύτητας. Δεν έχουν αντοχή σε υψηλές θερμοκρασίες και είναι απαραίτητη η χρήση ψυκτικού υγρού.
- **Ελαστικά υψηλής αντοχής ενισχυμένα με ίνες (Rubber reinforced)**
- **Μεταλλικά υλικά όπως Αργίλιο, κράματα χαλκού-ορειχάλκου, μπρούτζος και σπανιότερα χάλυβας:** Χρησιμοποιούνται κυρίως σε αδαμαντοτροχούς, λόγω της πολύ υψηλής αντοχής τους, όπως επίσης και στους τροχούς ηλεκτροχημικής λείανσης, λόγω της ηλεκτρικής τους αγωγιμότητας.

Κάθε υλικό σύνδεσης συμβολίζεται με συγκεκριμένο τρόπο, σύμφωνα με τον παρακάτω πίνακα.

Είδος συνδετικού υλικού	Συμβολισμός
Κεραμικά υλικά	V
Υδρύαλος	S
Συνθετικές Ρητίνες	B
Ρητινικό υλικό ενισχυμένο με ίνες	BF
Ελαστικά υλικά υψηλής αντοχής	R
Ελαστικά υλικά υψηλής αντοχής ενισχυμένα με ίνες	RF
Μεταλλικά υλικά	M

Πίνακας 2: Συμβολισμός συνδετικών υλικών [3]

#### 1.4.4 Σκληρότητα (grade) λειαντικού τροχού

Η σκληρότητα του λειαντικού τροχού καθορίζεται από την σκληρότητα του συνδετικού υλικού του τροχού. Μαλακοί τροχοί χάνουν κόκκους εύκολα και χρησιμοποιούνται για την λείανση σκληρών υλικών όπου αναμένονται χαμηλοί ρυθμοί απώλειας λειαινόμενου υλικού. Αντίθετα, σκληροί τροχοί χρησιμοποιούνται για λείανση πιο μαλακών επιφανειών.

Η σκληρότητα των λειαντικών τροχών εκφράζεται διεθνώς με γράμματα της λατινικής αλφαβήτου, όπως φαίνεται στον παρακάτω πίνακα.

Χαρακτηρισμός τροχού	Σκληρότητα
Εξαιρετικά μαλακός	A, B, C, D
Πολύ μαλακός	E, F, G
Μαλακός	H, I, J, K
Μέσης σκληρότητας	L, M, N, O
Σκληρός	P, Q, R, S
Πολύ σκληρός	T, U, V, W
Εξαιρετικά σκληρός	X, Y, Z

Πίνακας 3: Κωδικοποίηση σκληρότητας λειαντικού τροχού [3]

#### 1.4.5 Υφή (structure) λειαντικού τροχού

Η υφή του τροχού δείχνει το μέγεθος των διάκενων μεταξύ των κόκκων σε αυτόν. Αραιή δομή σημαίνει περισσότεροι πόροι και λιγότεροι κόκκοι ανά μονάδα όγκου του τροχού. Η ανοικτή δομή συνιστάται για λείανση υλικών τα οποία έχουν την τάση να δημιουργούν συνεχόμενα απόβλητα, ενώ τροχοί με πυκνότερη υφή χρησιμοποιούνται για καλύτερο τελείωμα επιφάνειας και καλύτερη διαστασιακή ακρίβεια.

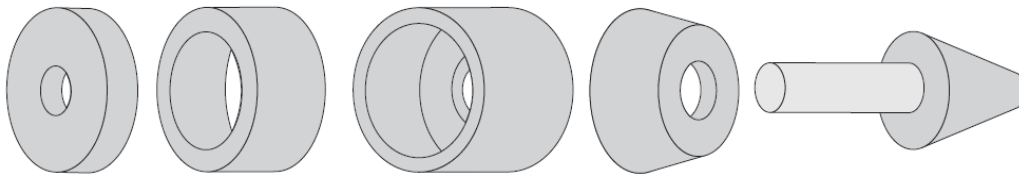
Η κατηγοριοποίηση των λειαντικών τροχών ανάλογα με την σκληρότητα εμφανίζεται στον παρακάτω πίνακα.

Χαρακτηρισμός τροχού	Υφή
Πολύ πυκνή υφή	0, 1
Πυκνή υφή	2, 3
Μέση υφή	4, 5
Ανοικτή υφή	6, 7
Πολύ ανοικτή υφή	8, 9

Πίνακας 4: Κωδικοποίηση υφής λειαντικού τροχού [3]

#### 1.4.6 Γεωμετρία τροχού λείανσης

Οι τροχοί λείανσης παράγονται σε μία μεγάλη ποικιλία από γεωμετρίες, οι συνηθέστεροι των οποίων παρουσιάζονται στο παρακάτω σχήμα:



Σχήμα 14: Πιθανές γεωμετρίες λειαντικών τροχών [1]

#### 1.4.7 Προδιαγραφές και κωδικοποίηση τροχών λείανσης

Οι λειαντικοί τροχοί κωδικοποιούνται με ένα σύστημα αριθμών και γραμμάτων της λατινικής αλφαβήτου οι οποίοι τον περιγράφουν. Η κωδικοποίηση αυτή κατά ISO παρουσιάζεται στο παρακάτω σχήμα [3]:

Σχήμα και διαστάσεις				Υλικό					
1	–	450 x 100 x 127		A	60	K	8	V	35
—	—	—	—	—	—	—	—	—	—
Βασική μορφή κατά ISO	Εξωτερική διάμετρος	Πλάτος	Διάμετρος οπής	Κοιτικό υλικό	Κόκκωση	Βαθμός σκληρότητας (μαλακός)	Χαρακτηριστικός αριθμός υφής	Συνδετικό υλικό (ν: κεραμικό)	Επιτρεπόμενη μέγιστη περιφερειακή ταχύτητα (m/s)

Σχήμα 15: Κωδικοποίηση λειαντικών τροχών κατά ISO [3]

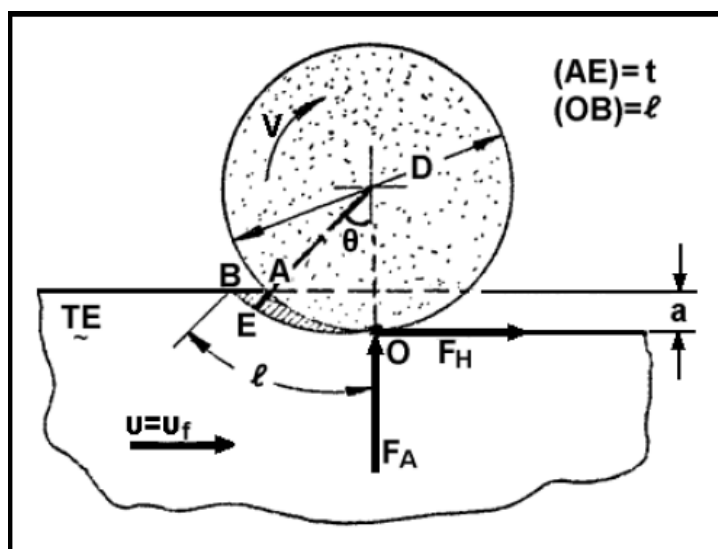
## 1.5 Μηχανική Λείανσης

Στην παρούσα παράγραφο επιχειρείται μία σύντομη παρουσίαση της βασικής μηχανικής για διάφορα είδη λείανσης

### 1.5.1 Επίπεδη περιφερειακή λείανση

Ο μηχανισμός της επίπεδης περιφερειακής λείανσης παρουσιάζεται στο παρακάτω σχήμα.





Σχήμα 16: Μηχανισμός επίπεδης περιφερειακής λείανσης [3]

Στο συγκεκριμένο είδος λείανσης, ισχύουν τα παρακάτω.

#### 1.5.1.1 Πάχος αποβλήτου

Τα βασικά χαρακτηριστικά του αποβλήτου, είναι τα παρακάτω, όπως φαίνονται στο σχήμα:

- Καμπύλη  $OA$ : Η διαδρομή του πρώτου λειαντικού κόκκου κατά την κίνηση του τεμαχίου
- Καμπύλη  $OB$ : Η διαδρομή του δεύτερου λειαντικού κόκκου κατά την σχετική κίνηση του τεμαχίου
- Ευθεία  $AB$ : Η πρόωση ανά κόκκο
- $AE$ : Το μέγιστο πάχος αποβλήτου ( $t$ )

Επίσης γίνονται οι εξής παραδοχές:

- Το μήκος του αποβλήτου ( $l$ ) είναι πολύ μεγαλύτερο από το μέγιστο πάχος του ( $t$ )
- Το μέσο πάχος του αποβλήτου  $\bar{t}$  ισούται με το μισό του μέγιστου πάχους, δηλ.  

$$\bar{t} = t/2$$
- Ο μέσος όγκος αποβλήτων  $Q$  ισούται με:

$$Q = l \bar{t} \bar{b}$$

όπου  $\bar{b}$  το μέσο πλάτος του αποβλήτου.

Επίσης ισχύει ότι:

$$Q = \frac{\text{Ρυθμός αποβολής υλικού}}{\text{Ρυθμός παραγωγής αποβλήτων}} = \frac{v b a}{V b C} = \frac{v a}{V C}$$

όπου  $v$  η ταχύτητα του τεμαχίου,  $V$  η ταχύτητα του λειαντικού τροχού,  $b$  το πλάτος επαφής του τροχού,  $a$  το βάθος λείανσης και  $C$  ο αριθμός κόκκων ανά μονάδα επιφάνειας του τροχού.

Συνδυάζοντας τις δύο παραπάνω σχέσεις, αποδεικνύεται ότι το πάχος του αποβλήτου δίνεται από την σχέση:

$$t = \frac{2 v a}{V C l b}$$

### 1.5.1.2 Μήκος επαφής

Με βάση την γεωμετρία της επίπεδης λείανσης, αποδεικνύεται [3] ότι το μήκος του αποβλήτου δίνεται από την σχέση:

$$l = \sqrt{aD}$$

### 1.5.1.3 Μέγιστο βάθος κοπής κόκκου

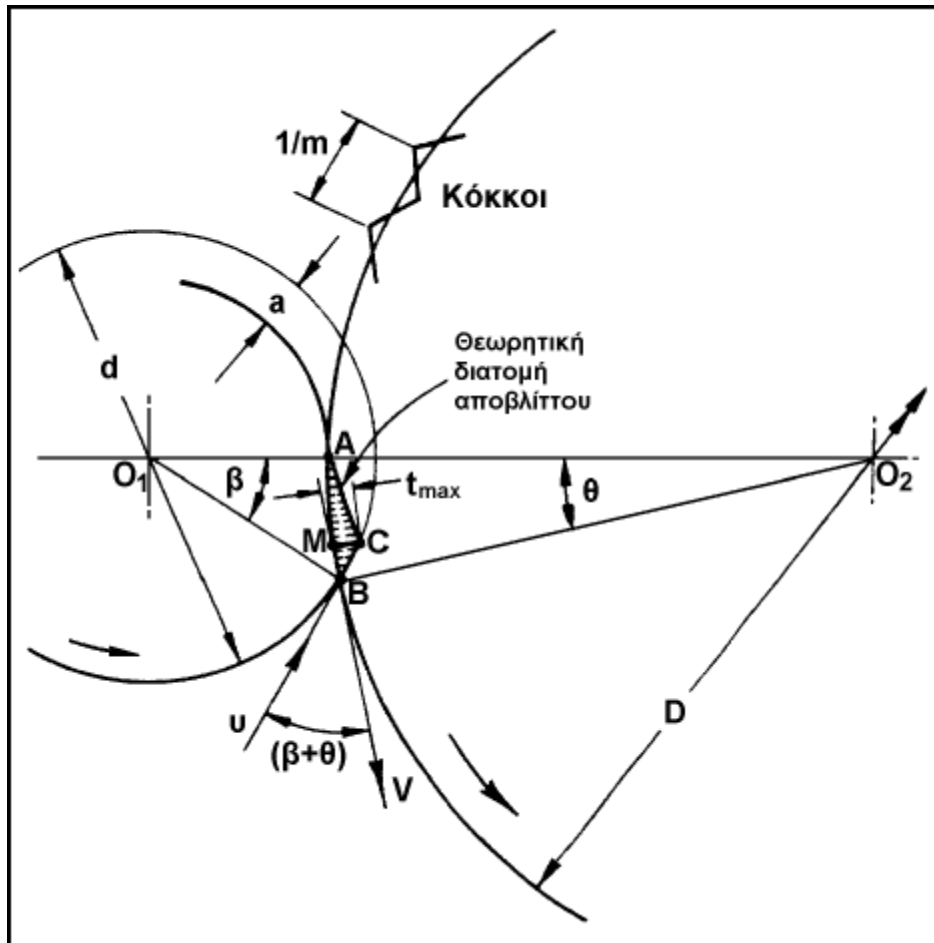
Έστω το αδιάστατο μέγεθος  $r = b/t$  το οποίο είναι μία χαρακτηριστική ποσότητα για την λείανση και είναι δυνατόν να μετρηθεί πειραματικά. Για τις περισσότερες εφαρμογές, ο λόγος αυτός λαμβάνει τιμές μεταξύ 5 και 20.

Αποδεικνύεται [3] ότι το μέγιστο πάχος του αποβλήτου κατά την λείανση δίνεται από την σχέση:

$$t = \sqrt{\frac{4 v}{C V r} \sqrt{a} D}$$

## 1.5.2 Κυλινδρική λείανση

Η γεωμετρία της κυλινδρικής περιφερειακής λείανσης παρουσιάζεται στο παρακάτω σχήμα:



Σχήμα 17: Γεωμετρία κυλινδρικής λείανσης [3]

Σε αυτό το είδος λείανσης, ισχύουν τα παρακάτω.

### 1.5.2.1 Μέγιστο βάθος κοπής ανά κόκκο

Έστω  $m$  ο αριθμός ενεργών κόκκων ανά μονάδα μήκους της περιφέρειας του τροχού. Η μέση απόσταση μεταξύ των δύο γειτονικών κόκκων είναι  $1/m$ , οπότε ο χρόνος που παρεμβάλλεται για την διέλευσή τους από τυχόν σημείο, θα είναι

$$\tau = \frac{1}{V m}$$

Σε αυτό το χρονικό διάστημα, το τεμάχιο και ο τροχός διανύουν τις εξής αποστάσεις αντιστοίχως:

$$(AB) = V \tau$$

$$(BC) = V \tau$$

Το μέγιστο πάχος αποβλήτου είναι ίσο με το ύψος ( $CM$ ) του τριγώνου  $ACB$ .

Αποδεικνύεται [3] ότι το μέγιστο βάθος κοπής δίνεται από την σχέση:

$$t = \frac{v}{mV} \sqrt{4a \left( \frac{D+d}{Dd} \right)}$$

Λαμβάνοντας υπ' όψη ότι ο αριθμός κόκκων ανά μονάδα μήκους  $m$  συνδέεται με τον αριθμό κόκκων ανά μονάδα επιφάνειας με την σχέση:

$$m = \frac{C r t}{2}$$

προκύπτει ότι:

$$t = \sqrt{\frac{4v}{C r V} \sqrt{a \left( \frac{D+d}{Dd} \right)}}$$

### 1.5.2.2 Ισοδύναμη διάμετρος

Θέτοντας στην τελευταία εξίσωση

$$D_e = \frac{Dd}{D+d}$$

προκύπτει:

$$t = \sqrt{\frac{4 v}{C r V} \sqrt{\frac{a}{D_e}}}$$

Παρατηρείται ότι η παραπάνω εξίσωση είναι δυνατόν να περιγράψει οποιαδήποτε είδος λείανσης.

Το μέγεθος  $D_e$  ονομάζεται ισοδύναμη διάμετρος. Το θετικό πρόσημο αντιστοιχεί σε εξωτερική κυλινδρική λείανση, το αρνητικό σε εσωτερική ενώ όταν το  $d$  είναι άπειρο, τότε  $D_e = D$  και η λείανση είναι επίπεδη.

### 5.2.3 Μήκος επαφής

Όπως και στην επίπεδη λείανση, αποδεικνύεται [3] ότι το μήκος επαφής δίνεται από την εξής σχέση:

$$l = \sqrt{aD_e}$$

### 1.5.3 Δυνάμεις και ισχύς λείανσης

Οι δυνάμεις  $F_H$  και  $F_A$  (βλ. Σχήμα 15) αντιστοιχούν στην κύρια δύναμη κοπής και στην δύναμη άπωσης οι οποίες αναπτύσσονται στον λειαντικό τροχό κατά την λείανση. Από αυτές, η  $F_A$  δεν παράγει έργο, καθώς είναι κάθετη στην κίνηση. Εκτός αυτού, έχει αποδειχθεί πειραματικά ότι κατά την λείανση, η δύναμη άπωσης είναι περίπου διπλάσια της δύναμης κοπής [3].

Η ενέργεια κοπής  $u$  ανά μονάδα όγκου αφαιρούμενου υλικού αποδεικνύεται [3] ότι στην επίπεδη λείανση είναι ίσο με:

$$u = \frac{F_H V}{v b a}$$

Η απαιτούμενη ισχύς λείανσης υπολογίζεται από την σχέση:

$$N = F_H V$$

#### 1.5.4 Δρώσα δύναμη ανά κόκκο

Η μέση δύναμη κοπής ανά κόκκο, αποδεικνύεται [3] ότι είναι ίση με:

$$F'_H = \frac{u r t^2}{4}$$

Η συμπεριφορά του λειαντικού τροχού εξαρτάται από το μέγεθος  $F'_H$ . Όσο μεγαλύτερες είναι οι δυνάμεις που δύναται να αναπτύξει ένας λειαντικός τροχός, τόσο μικρότερη είναι η πραγματική του σκληρότητα κατά την λείανση.

Συγκεκριμένα, από την παραπάνω εξίσωση είναι δυνατόν να συμπεράνουμε, ότι για να μειωθεί η πραγματική σκληρότητα του τροχού κατά την λείανση, χρειάζεται να συντελεστούν τα παρακάτω:

- Αύξηση του βάθους κοπής  $t$
- Αύξηση της ταχύτητας πρόωσης ( $v_f$  ή  $v$ )
- Μείωση της ταχύτητας λείανσης  $V$
- Αύξηση του βάθους λείανσης  $a$
- Αύξηση της καμπυλότητας της επιφάνειας τεμαχίου, δηλαδή μείωση της διαμέτρου  $d$
- Μείωση του αριθμού λειαντικών κόκκων ανά μονάδα επιφάνειας του λειαντικού τροχού

## 2. Νευρωνικά δίκτυα

### 2.1 Εισαγωγή

Η ιδέα των νευρωνικών δικτύων ξεκίνησε με την παρατήρηση και μελέτη των βασικών αρχών πάνω στις οποίες στηρίζεται η λειτουργία του ανθρώπινου εγκεφάλου [5]. Πρόκειται για ένα εξαιρετικά πολύπλοκο μη γραμμικό σύστημα, του οποίου τα βασικά δομικά στοιχεία είναι κάποιες στοιχειώδεις υπολογιστικές μονάδες γνωστές ως νευρώνες. Οι νευρώνες συνδέονται με ένα είδος συνάψεων μέσω των οποίων ανταλλάσσονται δεδομένα.

Ο ανθρώπινος εγκέφαλος είναι σε θέση να εκτελεί ορισμένους υπολογισμούς πολύ γρηγορότερα και από τον ισχυρότερο υπολογιστή, όπως π.χ. αναγνώριση εικόνας, αναγνώριση ήχου κ.τ.λ.. Τα αποτελέσματα αυτά οφείλονται στο γεγονός ότι είναι σε θέση να προσαρμόζει την δομή του σε νέα δεδομένα, αποθηκεύοντας πληροφορίες οι οποίες ονομάζονται εμπειρία. Η ταχύτητα του εγκεφάλου οφείλεται και στην παράλληλη σύνδεση μεταξύ των νευρώνων, πράγμα το οποίο ισχύει και για τα νευρωνικά δίκτυα.

Σε αναλογία με τον ανθρώπινο εγκέφαλο, ένα τεχνητό νευρωνικό δίκτυο είναι ένα σύστημα αποτελούμενο από νευρώνες και τις μεταξύ τους συνδέσεις το οποίο είναι κατασκευασμένο για να εκτελεί έναν συγκεκριμένο σκοπό, όπως π.χ. αναγνώριση εικόνας, ήχου ή προσέγγιση συνάρτησης και πειραματικών δεδομένων. Επίσης, όπως και ο εγκέφαλος, το τεχνητό νευρωνικό δίκτυο, δύναται να προσαρμόζει την δομή του σε νέα δεδομένα τα οποία εισάγονται σε αυτό. Η διαδικασία αυτή είναι γνωστή ως εκπαίδευση ή εκμάθηση και επιτυγχάνεται μέσω της εκτέλεσης συγκεκριμένου αλγορίθμου ο οποίος αναφέρεται σαν αλγόριθμος εκπαίδευσης (learning algorithm).

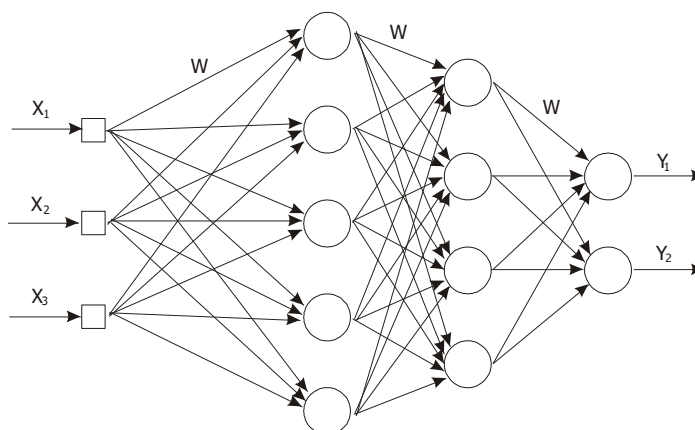
Τα νευρωνικά δίκτυα παρουσιάζουν συγκεκριμένα πλεονεκτήματα σε σύγκριση με τις συμβατικές τεχνικές τα οποία είναι εντοπισμένα στην δυνατότητα γενίκευσης (generalization) των παραγόμενων αποτελεσμάτων. Αυτό σημαίνει ότι το νευρωνικό δίκτυο παράγει σωστά αποτελέσματα ακόμη και για δεδομένα που είναι διαφορετικά από αυτά που χρησιμοποιήθηκαν κατά την διαδικασία εκτέλεσης του αλγορίθμου εκμάθησης, αρκεί οι τιμές τους να βρίσκονται μέσα στο εύρος που ορίζεται από τα δεδομένα εκπαίδευσης. Επίσης, δεν είναι απαραίτητη η χρήση του φυσικού

μηχανισμού που διέπει το εκάστοτε πρόβλημα, η οποία μπορεί να έχει απαγορευτική πολυπλοκότητα, ή απλώς να μην είναι γνωστή. Τα συγκεκριμένα πλεονεκτήματα, δίνουν την ικανότητα στα νευρωνικά δίκτυα να επιλύουν μεγάλης κλίμακας, μη γραμμικά και υψηλής πολυπλοκότητας προβλήματα.

## 2.2 Λειτουργία των νευρωνικών δικτύων

Όπως αναφέρθηκε και στην εισαγωγή, ένα νευρωνικό δίκτυο αποτελείται από ένα σύνολο νευρώνων οι οποίοι επικοινωνούν μεταξύ τους μέσω συνάψεων. Οι νευρώνες διατάσσονται σε παράλληλες σειρές οι οποίες ονομάζονται στοιβάδες [6]. Ένα νευρωνικό δίκτυο δύναται να έχει απεριόριστες στοιβάδες νευρώνων και απεριόριστους νευρώνες σε κάθε στοιβάδα. Ο μοναδικός περιορισμός είναι ο υπολογιστικός χρόνος κατά την διάρκεια της εκτέλεσης του αλγορίθμου εκπαίδευσης καθώς, όπως είναι προφανές, όσο περισσότερο πολύπλοκη είναι η δομή ενός νευρωνικού δικτύου τόσο δυσκολότερα εκτελείται ο τελευταίος. Μερικές από τις πιο συνηθισμένες μορφές νευρωνικών δικτύων που έχουν προταθεί περιγράφονται στις επόμενες παραγράφους.

Ένα τυπικό νευρωνικό δίκτυο τύπου πρόσθιας τροφοδότησης (feedforward neural network) το οποίο έχει πέντε νευρώνες στην πρώτη στοιβάδα, τέσσερις νευρώνες στην δεύτερη και δύο στην τρίτη εικονίζεται στο παρακάτω σχήμα. Η πρώτη και η δεύτερη στοιβάδα νευρώνων ονομάζονται κρυφές στοιβάδες (hidden layers) καθώς δεν παράγουν τελική έξοδο ούτε και δέχονται τις αρχικές τιμές εισόδου δηλαδή εκτελούν καθαρά εσωτερικές πράξεις του δικτύου. Το σχήμα παρουσιάζει ένα δίκτυο το οποίο αποτελεί ένα διάγραμμα ροής πληροφοριών.



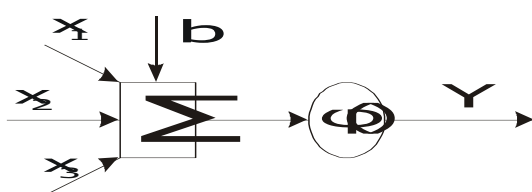
Σχήμα 18: Νευρωνικό δίκτυο με τρεις κόμβους εισαγωγής δεδομένων, πέντε νευρώνες στην πρώτη στοιβάδα, τέσσερις στην δεύτερη και δύο στην τρίτη



Τα δεδομένα εισόδου εισάγονται αρχικά στους κόμβους εισόδου. Κάθε κόμβος συνδέεται με όλους τους νευρώνες της πρώτης στοιβάδας και μεταφέρει σε αυτούς όλες τις πληροφορίες που λαμβάνει, χωρίς να εκτελέσει καμία μαθηματική πράξη. Σημειώνεται ότι οι πληροφορίες μεταφέρονται μόνον κατά την κατεύθυνση των βελών στις συνδέσεις. Κάθε νευρώνας εκτελεί μία μαθηματική πράξη η οποία θα αναλυθεί στην συνέχεια και το αποτέλεσμα αυτής μεταφέρεται μέσω των συνάψεων σε όλους τους νευρώνες της δεύτερης στοιβάδας. Οι νευρώνες της δεύτερης στοιβάδας εκτελούν και αυτοί μαθηματικές πράξεις και μεταφέρουν τα αποτελέσματά τους στους νευρώνες της τρίτης οι οποίοι παράγουν την τελική έξοδο του νευρωνικού δικτύου.

Κάθε σύναψη η οποία μεταφέρει μία πληροφορία είναι εφοδιασμένη με έναν συντελεστή ο οποίος ονομάζεται βάρος και συμβολίζεται με  $w$ . Η πληροφορία η οποία συνίσταται ουσιαστικά από έναν αριθμό και μεταφέρεται μέσω της σύναψης, πολλαπλασιάζεται με το βάρος  $w$  προτού εισαχθεί στον νευρώνα του επομένου στρώματος. Σημειώνεται ότι τα βάρη είναι γενικά διαφορετικά για κάθε σύναψη.

Η δομή ενός νευρώνα παρουσιάζεται στο σχήμα 19. Ο νευρώνας έχει τρεις μεταβλητές εισόδου τις  $x_1$ ,  $x_2$  και  $x_3$ , οι οποίες έχουν ήδη πολλαπλασιαστεί με τα βάρη  $w$  και στη συνέχεια αθροίζει προσθέτοντας ένα σταθερό αριθμό  $b$ , χαρακτηριστικό για κάθε νευρώνα [7]. Στη συνέχεια εφαρμόζει στο άθροισμα μία συνάρτηση, για να υπολογίσει το τελικό αποτέλεσμα.



**Σχήμα 19:** Δομή ενός νευρώνα του νευρωνικού δικτύου

Μία τυπική μη γραμμική συνάρτηση που είναι δυνατόν να χρησιμοποιηθεί για τον σκοπό αυτόν είναι η συνάρτηση υπερβολικής εφαπτομένης που ορίζεται από την παρακάτω σχέση.

$$\phi(x) = \tanh(x)$$

Προκειμένου το δίκτυο να εκτελεί την εργασία για την οποία είναι προορισμένο, είναι απαραίτητο κάθε νευρώνας να παράγει τις επιθυμητές εξόδους όταν του δοθούν νέες εισόδοι. Για να γίνει αυτό απαιτείται ο προσδιορισμός των παραμέτρων του νευρωνικού δικτύου, ήτοι των βαρών  $w$  και των παραμέτρων  $b$  οι οποίες προστίθενται στο άθροισμα των εισόδων.

Αυτό επιτυγχάνεται δίνοντας στο νευρωνικό δίκτυο έναν αριθμό έτοιμων δεδομένων εισόδου και εξόδου του εξεταζόμενου προβλήματος έτσι ώστε το δίκτυο να προσαρμόσει τις παραμέτρους του σε αυτά. Τα δεδομένα αυτά ονομάζονται δεδομένα εκπαίδευσης και η διαδικασία προσαρμογής των παραμέτρων σε αυτά ονομάζεται αλγόριθμος εκπαίδευσης.

Προτού εκτελεστεί ο αλγόριθμος εκπαίδευσης είναι απαραίτητη η επιλογή τόσο της δομής του νευρωνικού δικτύου, δηλαδή του αριθμού των στοιβάδων και των νευρώνων που περιέχονται σε αυτές, αλλά και η επιλογή της κατάλληλης συνάρτησης στους νευρώνες. Η επιλογή του αριθμού των στοιβάδων και των νευρώνων είναι η ουσιαστικότερη και δυσκολότερη προς προσδιορισμό παράμετρος. Δεν υπάρχει συγκεκριμένος αλγόριθμος υπολογισμού οπότε επιλέγεται εμπειρικά από τον χρήστη. Εάν ο αριθμός των νευρώνων δεν είναι επαρκής τότε το δίκτυο δεν είναι σε θέση να παρακολουθήσει όλες τις διακυμάνσεις της εξόδου στα δεδομένα εκπαίδευσης οπότε εκ των πραγμάτων δεν θα είναι σε θέση να κάνει ακριβείς προβλέψεις. Εάν αντίθετα είναι υπερβολικά μεγάλος τότε παρατηρείται σε πολλές περιπτώσεις ένα ανεπιθύμητο φαινόμενο, κατά το οποίο το δίκτυο προσαρμόζει τις παραμέτρους του όχι μόνο στα πραγματικά δεδομένα αλλά και στον θόρυβο που εμπεριέχεται σε αυτά και κατά συνέπεια η ικανότητα πρόβλεψης στα καινούργια δεδομένα ελαττώνεται. Το φαινόμενο αυτό ονομάζεται *overfitting*.

## **2.3 Αρχιτεκτονικές των νευρωνικών δικτύων**

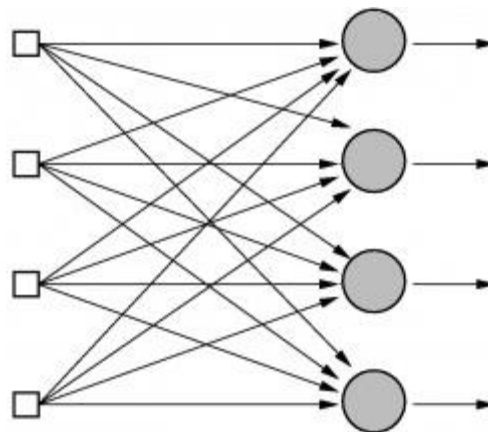
Ανάλογα με το πρόβλημα που πρέπει να επιλυθεί χρησιμοποιείται και το κατάλληλο είδος νευρωνικών δικτύων. Παρ' όλο που υπάρχουν διάφορα είδη, όλα τα νευρωνικά δίκτυα μοιράζονται την ίδια βασική δομή, δηλαδή αποτελούνται από

νευρώνες και παράλληλες συνδέσεις μέσω των οποίων ανταλλάσσονται πληροφορίες..

Η αρχιτεκτονική ενός νευρωνικού δικτύου καθορίζεται από τον αριθμό των νευρώνων και την διάταξη των συνδέσεων μεταξύ αυτών. Η αρχιτεκτονική καθορίζει τον τρόπο λειτουργίας του νευρωνικού δικτύου, καθώς επίσης και την επιλογή του αλγορίθμου εκπαίδευσης. Στην ακόλουθη παράγραφο αναλύονται οι τρεις τυπικές διαφορετικές διατάξεις νευρωνικών δικτύων που έχουν προταθεί.

### 2.3.1 Δίκτυα πρόσθιας τροφοδότησης ενός επιπέδου

Πρόκειται για την απλούστερη μορφή νευρωνικού δικτύου, όπου οι νευρώνες οργανώνονται σε επίπεδα. Το πρώτο επίπεδο είναι το επίπεδο εισόδου των δεδομένων, το οποίο στέλνει δεδομένα στο δεύτερο επίπεδο που αποτελεί το επίπεδο υπολογισμού και εξόδου ταυτόχρονα. Τυπική μορφή ενός τέτοιου νευρωνικού δικτύου απεικονίζεται στο σχήμα 20 [8].



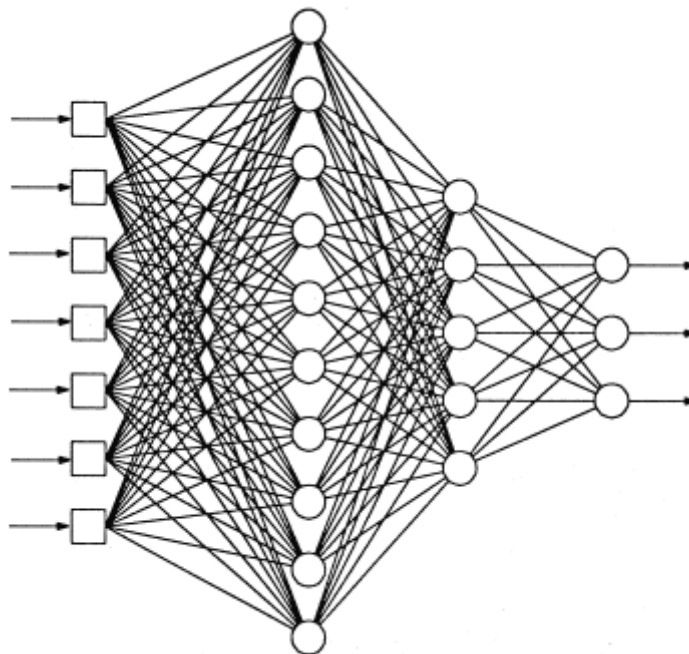
Σχήμα 20: Δίκτυο πρόσθιας τροφοδότησης ενός επιπέδου

Ο προσδιορισμός ενός επιπέδου αντιστοιχεί στους νευρώνες όπου εκτελούνται υπολογισμοί. Οι νευρώνες εισόδου δεδομένων, καθώς δεν εκτελούν υπολογισμούς δεν προσμετρώνται στον αριθμό των επιπέδων.

### 2.3.2 Δίκτυα πρόσθιας τροφοδότησης πολλαπλών επιπέδων

Η κατηγορία αυτή των νευρωνικών δικτύων χαρακτηρίζεται από την ύπαρξη ενός ή περισσότερων κρυμμένων επιπέδων νευρώνων, οι οποίοι παρεμβάλλονται μεταξύ του επιπέδου νευρώνων εισαγωγής και εξόδου του δικτύου. Ένα τυπικό

νευρωνικό δίκτυο με 7 νευρώνες εισόδου, 3 νευρώνες εξόδου και δύο κρυμμένα επίπεδα νευρώνων με 10 και 5 νευρώνες αντίστοιχα το καθένα, απεικονίζεται στο σχήμα 21.



Σχήμα 21: Νευρωνικό δίκτυο πρόσθιας τροφοδότησης πολλαπλών επιπέδων [8]

Τα δεδομένα εισάγονται στο νευρωνικό δίκτυο από το πρώτο επίπεδο (κόμβοι πηγής) ως διάνυσμα εισόδου και χωρίς μετατροπή μεταφέρονται στο πρώτο κρυμμένο επίπεδο νευρώνων (κόμβοι υπολογισμών). Τα σήματα εξόδου από το πρώτο κρυμμένο επίπεδο, χρησιμοποιούνται σαν εισοδοί στο δεύτερο κρυμμένο επίπεδο νευρώνων και τέλος, τα σήματα εξόδου του δεύτερου κρυμμένου επιπέδου, χρησιμοποιούνται σαν εισοδοί στους νευρώνες εξόδου, απ' όπου και προέρχεται η τελική έξοδος του νευρωνικού δικτύου.

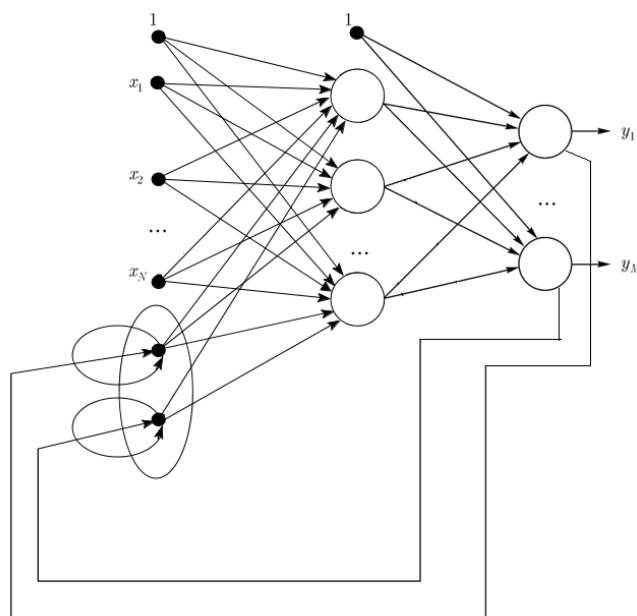
Με την προσθήκη περισσότερων επιπέδων, το νευρωνικό δίκτυο είναι ικανό να εκτελεί στατιστικούς υπολογισμούς υψηλότερης τάξης. Το δίκτυο στο σχήμα 21, αναφέρεται για συντομία, ως δίκτυο 7-10-5-3, δηλ, 7 κόμβοι πηγής, δύο κρυμμένα επίπεδα νευρώνων με 10 και 5 νευρώνες αντίστοιχα και 3 νευρώνες εξόδου.

Τονίζεται εδώ ότι η ροή των σημάτων είναι προς μία κατεύθυνση, δηλ, π.χ. από το πρώτο κρυμμένο επίπεδο, στο δεύτερο, από το δεύτερο στο τρίτο κ.ο.κ. Επίσης, το δίκτυο που εμφανίζεται στο σχήμα 21 είναι πλήρως συνδεδεμένο, δηλαδή κάθε

νευρώνας ενός επιπέδου συνδέεται με όλους τους νευρώνες του επομένου επιπέδου. Εφόσον έλειπε ένας αριθμός συνδέσεων, το δίκτυο θα ονομαζόταν μερικώς συνδεδεμένο.

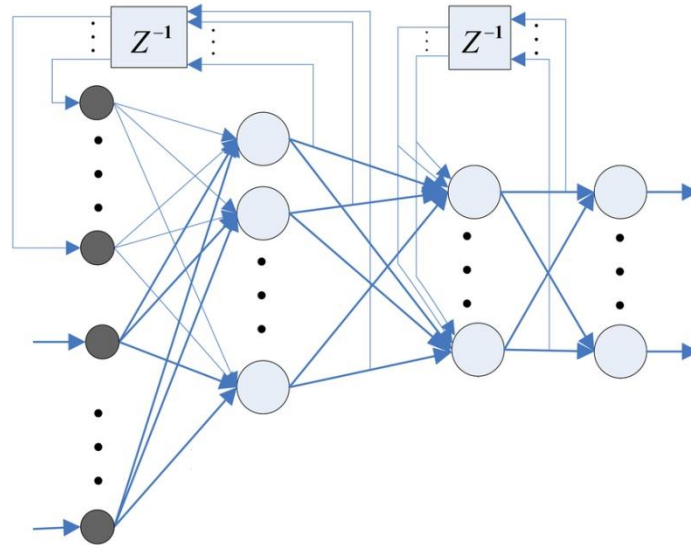
### 2.3.3 Δίκτυα με ανάδραση

Στις δύο προηγούμενες περιπτώσεις νευρωνικών δικτύων που αναλύθηκαν, η ροή πληροφορίας ήταν προς μία κατεύθυνση, από την είσοδο του δικτύου προς την έξοδο. Αντίθετα, ένα νευρωνικό δίκτυο με ανάδραση είναι ένα νευρωνικό δίκτυο πρόσθιας τροφοδότησης, το οποίο έχει τουλάχιστον έναν βρόχο ανατροφοδότησης. Ένα τέτοιο νευρωνικό δίκτυο απεικονίζεται στο σχήμα 22.



Σχήμα 22: Νευρωνικό δίκτυο με ανάδραση

Στο σχήμα 6, οι συνδέσεις ανατροφοδότησης, δημιουργούνται από την σύνδεση των νευρώνων εξόδου με τους νευρώνες εισόδου του δικτύου. Η παρουσία βρόγχων ανατροφοδότησης έχει σημαντικό αντίκτυπο στην απόδοση και στην δυνατότητα εκμάθησης του δικτύου [8]. Επίσης, οι βρόγχοι ανατροφοδότησης είναι δυνατόν να περιλαμβάνουν μονάδες καθυστέρησης, όπως φαίνεται στο σχήμα 23.



Σχήμα 23: Νευρωνικό δίκτυο με ανάδραση και μονάδες καθυστέρησης

## 2.4 Μορφές της συνάρτησης ενεργοποίησης των νευρωνικών δικτύων

Η συνάρτηση ενεργοποίησης, η οποία συμβολίζεται  $\varphi(v)$  καθορίζει την έξοδο ενός νευρώνα, με βάση το διάνυσμα εισόδου. Στις επόμενες παραγράφους αναλύονται οι διάφοροι τύποι συνάρτησης ενεργοποίησης που χρησιμοποιούνται στα νευρωνικά δίκτυα [8], [9].

### 2.4.1 Βηματική συνάρτηση (Heaviside function)

Ο συγκεκριμένος τύπος συνάρτησης ενεργοποίησης, ορίζεται από την παρακάτω εξίσωση:

$$\varphi(v) = \begin{cases} 1, & v \geq 0 \\ 0, & v < 0 \end{cases}$$

Κατά συνέπεια, με βάση τον παραπάνω ορισμό, η έξοδος του νευρώνα στην περίπτωση αυτήν θα δίνεται από την εξίσωση:

$$y_k = \begin{cases} 1, & v_k \geq 0 \\ 0, & v_k < 0 \end{cases}$$

όπου  $v_k$  είναι η είσοδος του νευρώνα και με την σειρά της ορίζεται από την εξίσωση:

$$v_k = \sum_{i=1}^n w_{ki}x_i + b_k$$

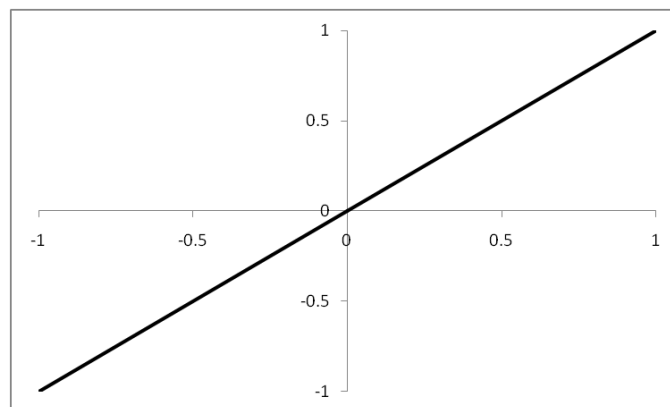
Ένας τέτοιος νευρώνας αναφέρεται στην βιβλιογραφία ως μοντέλο McCulloch - Pitts, σε αναγνώριση της εργασίας των δύο συγκεκριμένων ερευνητών, το 1943. Με την συγκεκριμένη συνάρτηση, η έξοδος του νευρώνα είναι μηδενική αν η είσοδος είναι αρνητική και 1 αν η είσοδος είναι θετική.

### 2.4.2 Γραμμική συνάρτηση

Η συγκεκριμένη συνάρτηση ενεργοποίησης, ορίζεται από την παρακάτω εξίσωση:

$$\varphi(v) = v$$

Η συγκεκριμένη συνάρτηση χρησιμοποιείται όταν οι νευρώνες χρειάζεται να εκτελούν μόνον γραμμικούς υπολογισμούς. Επίσης, χρησιμοποιείται στους νευρώνες εξόδου, όπου είναι απαραίτητο, η έξοδος να μην περιορίζεται στις τιμές  $[-1,1]$ . Η γραφική της παράσταση δίνεται στο σχήμα 24.



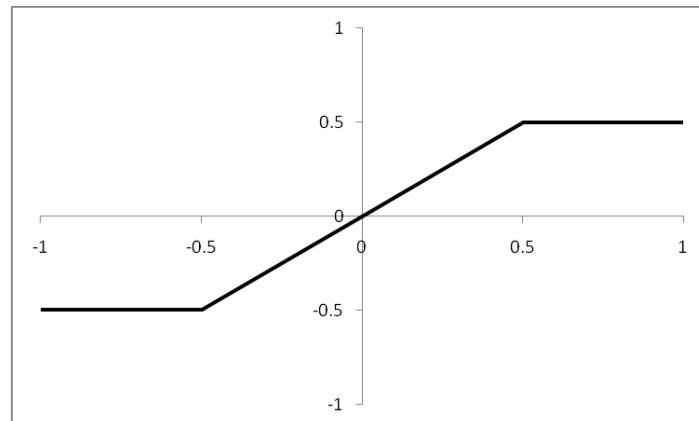
Σχήμα 24: Γραμμική συνάρτηση ενεργοποίησης

### 2.4.3 Γραμμική συνάρτηση ξεχωριστών βημάτων

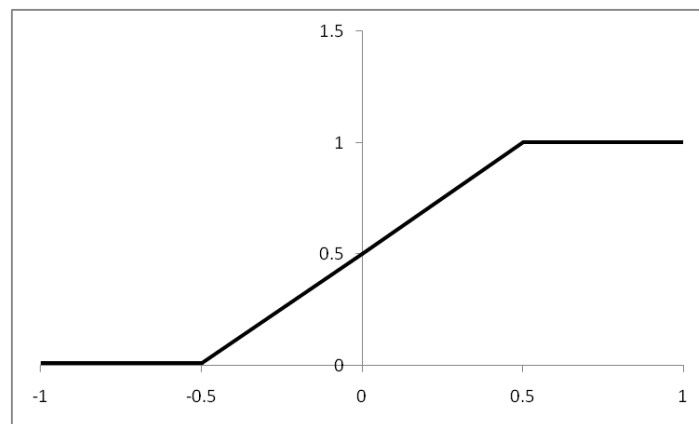
Η συνάρτηση αυτή, ορίζεται από την παρακάτω εξίσωση:

$$\varphi(v) = \begin{cases} 0, & v \leq -0.5 \\ v, & -0.5 \leq v \leq 0.5 \\ 1, & v \geq 0.5 \end{cases}$$

Η γραφική παράσταση της παραπάνω συνάρτησης δίδεται στο σχήμα 25 στην περίπτωση που δεν υπάρχει παράγοντας ενίσχυσης (bias) και στο σχήμα 26, στην περίπτωση όπου ο παράγοντας ενίσχυσης είναι ίσος με 0.5.



Σχήμα 25: Γραμμική συνάρτηση ξεχωριστών βημάτων χωρίς παράγοντα ενίσχυσης (bias)



Σχήμα 26: Γραμμική συνάρτηση ξεχωριστών βημάτων με bias = 0.5

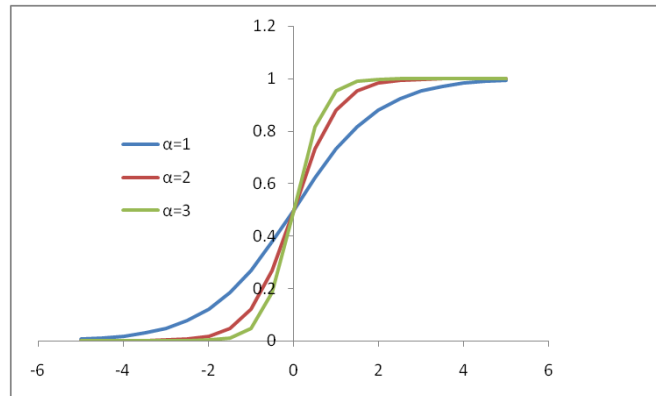
#### 2.4.4 Σιγμοειδής συνάρτηση

Η σιγμοειδής συνάρτηση είναι η συνηθέστερη μορφή συνάρτησης ενεργοποίησης για νευρωνικά δίκτυα. Είναι αυστηρώς αύξουσα συνάρτηση και παρουσιάζει μία ιδανική ισορροπία μεταξύ γραμμικής και μη γραμμικής συμπεριφοράς. Ένα παράδειγμα σιγμοειδούς συνάρτησης είναι η λογιστική συνάρτηση, η οποία δίδεται από την παρακάτω εξίσωση:



$$\varphi(v) = \frac{1}{1 + \exp(-\alpha v)}$$

όπου το  $\alpha$  είναι η παράμετρος κλίσεως της συνάρτησης αυτήν. Η επίδραση της παραμέτρου αυτής για διάφορες τιμές του  $\alpha$ , φαίνεται καθαρά στο σχήμα 27.

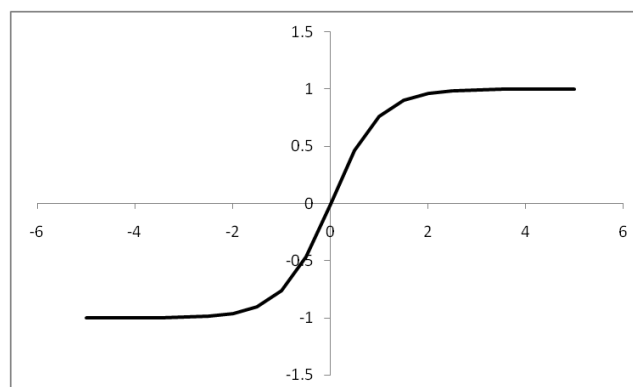


Σχήμα 27: Λογιστική συνάρτηση για διάφορες τιμές της παραμέτρου κλίσεως  $\alpha$

Μία άλλη μορφή σιγμοειδούς συναρτήσεως, η οποία χρησιμοποιείται ευρύτατα στα νευρωνικά δίκτυα είναι η συνάρτηση της υπερβολικής εφαπτομένης:

$$\varphi(v) = \tanh(v)$$

Το γράφημα της παραπάνω συνάρτησης απεικονίζεται στο σχήμα 28.



Σχήμα 28: Συνάρτηση υπερβολικής εφαπτομένης

Οι σιγμοειδής συναρτήσεις επιτρέπουν στις εξόδους των νευρώνων να πάρουν αρνητικές τιμές, το οποίο παρουσιάζει μία σειρά πλεονεκτημάτων.

## 2.5 Εκπαίδευση Νευρωνικών Δικτύων

Μετά τον καθορισμό του είδους νευρωνικού δικτύου που θα χρησιμοποιηθεί, τον αριθμό των κρυμμένων στοιβάδων νευρώνων, και τον αριθμό των νευρώνων σε κάθε στοιβάδα, σειρά έχει η εκπαίδευση του δικτύου. Με τον όρο εκπαίδευση εννοείται ο υπολογισμός των βαρών  $w_{ki}$  ώστε να ελαχιστοποιηθεί το σφάλμα πρόβλεψης του δικτύου. Οι μέθοδοι εκπαίδευσης είναι δυνατόν να χωριστούν σε δύο μεγάλες κατηγορίες [8], [9]:

- **Μη επιτηρούμενη εκπαίδευση νευρωνικών δικτύων** ονομάζεται η εκπαίδευση, η οποία πραγματοποιείται μόνον με διανύσματα εισόδου. Η συγκεκριμένη κατηγορία αλγορίθμων εκπαίδευσης εφαρμόζεται, όταν ένα νευρωνικό δίκτυο χρησιμοποιείται για να εντάξει μία είσοδο σε κάποια κατηγορία (clustering). Τα νευρωνικά δίκτυα τα οποία υφίστανται μη επιτηρούμενη εκπαίδευση έχουν την ικανότητα να αναγνωρίζουν πρότυπα με ατελή δεδομένα.
- **Επιτηρούμενη εκπαίδευση νευρωνικών δικτύων** ονομάζεται η εκπαίδευση κατά την οποία κάθε διάνυσμα εισόδου αντιστοιχίζεται με ένα διάνυσμα εξόδου, το οποίο αποτελεί την επιθυμητή έξοδο του νευρωνικού δικτύου. Η επιτηρούμενη εκπαίδευση νευρωνικών δικτύων πραγματοποιείται σε δύο στάδια. Στο πρώτο στάδιο δίνεται μία ομάδα διανυσμάτων εισόδου-εξόδου, τα οποία χρησιμοποιούνται για να υπολογιστούν τα βάρη και τα bias. Στο δεύτερο στάδιο, και αφού έχουν υπολογιστεί τα βάρη και τα bias, δίδεται στο νευρωνικό δίκτυο μία δεύτερη ομάδα δεδομένων εισόδου-εξόδου, τα οποία δεν χρησιμοποιήθηκαν κατά την εκπαίδευση, για να ελεγχθεί εάν το νευρωνικό δίκτυο ανταποκρίνεται σωστά σε νέα δεδομένα (validation).

Καθώς το αντικείμενο της παρούσας εργασίας είναι η μαθηματική μοντελοποίηση μίας φυσικής κατεργασίας με νευρωνικά δίκτυα, είναι δεδομένο ότι οι μέθοδοι εκπαίδευσης που χρησιμοποιούνται, αντιστοιχούν στην δεύτερη κατηγορία. Συγκεκριμένα, είναι διαθέσιμος ένας αριθμός πειραματικών δεδομένων εισόδου και εξόδου και δοκιμάζονται διαφορετικά νευρωνικά δίκτυα και αλγόριθμοι εκπαίδευσης, με σκοπό την μαθηματική μοντελοποίηση του συγκεκριμένου συστήματος.

## 2.5.1 Νόμοι εκπαίδευσης

Όπως αναφέρθηκε στην προηγούμενη παράγραφο, επιτηρούμενη εκπαίδευση ενός νευρωνικού δικτύου είναι διαδικασία κατά την οποία υπολογίζονται τα βάρη και τα bias του δικτύου, ώστε μετά την ολοκλήρωση της εκπαίδευσης, το νευρωνικό δίκτυο να είναι σε θέση να κάνει προβλέψεις βασισμένες σε πειραματικά δεδομένα.

Όλοι οι αλγόριθμοι εκπαίδευσης αποσκοπούν στον υπολογισμό των βαρών και των bias των νευρωνικών δικτύων και ακολουθούν μία σειρά νόμων, μερικοί εκ των οποίων αναφέρονται εν συντομία παρακάτω [9]:

- **Νόμος του Hebb:** Πρόκειται για τον γνωστότερο κανόνα εκπαίδευσης, ο οποίος προέρχεται από την μελέτη των νευρώνων του ανθρώπινου εγκεφάλου, και προτάθηκε από τον Hebb στο βιβλίο *The Organization of Behavior*, το 1949. Σύμφωνα με τον κανόνα, αν η σύνδεση δύο νευρώνων είναι πολύ ενεργή, τότε επέρχονται οι απαραίτητες μεταβολές, ώστε να αυξάνεται η αποδοτικότητα της σύνδεσης αυτής. Αυτό σε όρους τεχνητών νευρωνικών δικτύων μεταφράζεται ως εξής: Εάν δύο νευρώνες με σύνδεση μεταξύ τους έχουν το ίδιο πρόσημο στις εξόδους τους, τότε η τιμή του συντελεστή βαρύτητας μεταξύ τους είναι απαραίτητο να αυξηθεί.
- **Νόμος του Hopfield:** Ο νόμος του Hopfield είναι μία επέκταση του κανόνα του Hebb, ποσοτικοποιώντας την αύξηση ή μείωση των βαρών και των bias του νευρωνικού δικτύου. Σύμφωνα με τον κανόνα αυτόν, εάν ο συντελεστής βαρύτητας αυξάνεται, ο βαθμός αύξησης ή μείωσης ενός βάρους  $w_{kj}$  θα πρέπει να είναι ανάλογος με τον ρυθμό εκπαίδευσης του δικτύου.
- **Κανόνας Δέλτα:** Σε έναν αλγόριθμο εκπαίδευσης νευρωνικού δικτύου, ο κανόνας δέλτα είναι ένας κανόνας ελαχιστοποίησης μέσω παραγώγου (gradient descent), με σκοπό την προσαρμογή των βαρών και των bias των νευρωνικών δικτύων. Για έναν νευρώνα  $m$  με συνάρτηση ενεργοποίησης  $\varphi(v)$ , ο κανόνας δέλτα για την προσαρμογή του βάρους  $w_{lm}$  δίνεται από την παρακάτω εξίσωση:

$$\Delta w_{lm} = \alpha(d_m - y_m)\varphi'(h_m)x_l$$

όπου  $a$  είναι μία σταθερά η οποία ονομάζεται ρυθμός εκμάθησης,  $h_m$  είναι το άθροισμα των εισόδων του νευρώνα  $m$  και  $d_m$  είναι η επιθυμητή έξοδος του νευρώνα  $m$ .

- Κανόνας εκπαίδευσης του Kohonen: Εκτελείται στις μη επιτηρούμενες μεθόδους εκπαίδευσης και προέρχεται από τα βιολογικά συστήματα. Σύμφωνα με τον κανόνα αυτόν, οι νευρώνες ανταγωνίζονται μεταξύ τους, προκειμένου να ενημερωθούν οι συντελεστές βαρύτητάς τους.

## 2.5.2 Αλγόριθμοι εκπαίδευσης

Όπως προαναφέρθηκε, για την επιτηρούμενη εκπαίδευση νευρωνικού δικτύου χρησιμοποιείται μία ομάδα δεδομένων εισόδου-επιθυμητής εξόδου. Έστω ότι υπάρχουν  $Q$  ζευγάρια διανυσμάτων εισόδου/εξόδου  $x_q, d_q$ , όπου  $q = 1, 2, 3, \dots, Q$ , ο αριθμός του διανύσματος εισόδου εξόδου. Έστω επίσης  $o_{qi}$ , οι συνιστώσες του διανύσματος εξόδου του νευρωνικού δικτύου και  $d_{qj}$  οι επιθυμητές τιμές με βάση τις συγκεκριμένες εισόδους.

Πρώτο βήμα κατά την εκπαίδευση ενός νευρωνικού δικτύου είναι η εισαγωγή τυχαίων τιμών για τα βάρη  $w_{lm}$  και τα bias  $b_m$ . Ακολουθώντας, με βάση τα παραπάνω βάρη, για κάθε διάνυσμα εισόδου  $x_q$  υπολογίζεται η αντίστοιχη έξοδος του δικτύου  $o_q$ . Με βάση την έξοδο αυτήν, υπολογίζεται η απόκλιση από την επιθυμητή έξοδο  $d_q$  από την εξής σχέση [9]:

$$E_q = \frac{1}{2} \sum_{n=1}^N (d_{qn} - o_{qn})^2$$

Όπου  $N$  είναι ο αριθμός των νευρώνων της στοιβάδας εξόδου. Με βάση την παραπάνω σχέση, και κριτήριο την εύρεση του ελαχίστου της παραγώγου του σφάλματος, προσαρμόζονται τα βάρη και τα bias του νευρωνικού δικτύου. Παρ' όλο που υπάρχει ένας μεγάλος αριθμός τεχνικών και αλγορίθμων εκπαίδευσης, οι περισσότερες από αυτές στηρίζονται στην ελαχιστοποίηση της παραγώγου.

Η παραπάνω διαδικασία εφαρμόζεται για κάθε ζεύγος διανυσμάτων εισόδου-εξόδου. Όταν ολοκληρωθεί ολόκληρο το πέρασμα από όλα τα ζεύγη εισόδου/εξόδου, τότε συμπληρώνεται μία εποχή (epoch) εκπαίδευσης.

Δύο από τους συνηθέστερους και ευρύτερα χρησιμοποιούμενους αλγόριθμους εκπαίδευσης αναλύονται στις επόμενες παραγράφους.

### 2.5.2.1 Αλγόριθμος της οπισθοδρομικής διάδοσης (*back propagation*)

Ο αλγόριθμος της οπισθοδρομικής διάδοσης ανήκει στην κατηγορία της επιτηρούμενης εκπαίδευσης νευρωνικών δικτύων, εφαρμόζεται σε δίκτυα πρόσθιας τροφοδοσίας, τα οποία χρησιμοποιούν την σιγμοειδή συνάρτηση. Κατά την μέθοδο αυτή, αφού υπολογιστεί η τιμή του μέσου τετραγωνικού σφάλματος, διαδίδεται προς τα πίσω με στόχο την ελαχιστοποίησή του. Αυτή επιτυγχάνεται με μεταβολή των συντελεστών βαρύτητας και των bias προς την κατεύθυνση της αντίθετης τιμής της κλίσης της συνάρτησης ενεργοποίησης [9].

Υπάρχουν δύο εκδοχές του αλγορίθμου εκπαίδευσης της οπισθοδρομικής διάδοσης, οι οποίες θα παρουσιαστούν παρακάτω για ένα νευρωνικό δίκτυο με μία κρυμμένη στοιβάδα νευρώνων, όπου υπάρχουν  $L$  νευρώνες εισόδου αριθμημένοι σαν  $l = 1, 2, 3, \dots, L$ ,  $M$  νευρώνες στην κρυμμένη στοιβάδα αριθμημένοι σαν  $m = 1, 2, 3, \dots, M$  και  $N$  νευρώνες στην στοιβάδα εξόδου αριθμημένοι  $n = 1, 2, 3, \dots, N$ .

- **Περίπτωση 1: Οι τιμές των βαρών και των bias μεταβάλλονται μετά από εισαγωγή στο δίκτυο κάθε ζεύγους διανυσμάτων εισόδου/εξόδου.** Στην περίπτωση αυτή, τα βάρη προσαρμόζονται σύμφωνα με τις εξής σχέσεις:

$$w_{lm}^{new} = w_{lm}^{old} + \Delta w_{lm}$$

$$v_{mn}^{new} = v_{mn}^{old} + \Delta v_{mn}$$

Όπου με  $w_{lm}$  συμβολίζονται τα βάρη που αντιστοιχούν στην είσοδο των νευρώνων της κρυμμένης στοιβάδας και  $v_{mn}$  συμβολίζονται τα βάρη που αντιστοιχούν στους νευρώνες εξόδου. Επίσης,  $\Delta w_{lm}$  και  $\Delta v_{lm}$  είναι η μεταβολή των βαρών μετά από την εισαγωγή στο δίκτυο ενός καινούργιου ζεύγους διανυσμάτων εισόδου εξόδου. Η μεταβολή αυτή, είναι προς την κατεύθυνση της αρνητικής μερικής παραγώγου του μέσου τετραγωνικού σφάλματος απόκρισης του δικτύου, ήτοι:

$$\Delta w_{lm} = -a \frac{\partial E_q}{\partial w_{lm}}$$

$$\Delta v_{mn} = -a \frac{\partial E_q}{\partial v_{mn}}$$

όπου:

$E_q$ : Το μέσο τετραγωνικό σφάλμα της απόκρισης του νευρωνικού δικτύου για το διάνυσμα  $q$  και

$a$ : Ο ρυθμός εκπαίδευσης του νευρωνικού δικτύου ο οποίος κυμαίνεται μεταξύ 0 και 1. Μικρές τιμές ρυθμού εκπαίδευσης θα οδηγήσουν σε αργή σύγκλιση του δικτύου, ενώ αντίθετα μεγάλες τιμές ρυθμού εκπαίδευσης θα οδηγήσουν σε αριθμητική αστάθεια και απόκλιση των υπολογισμών.

Έχοντας δεδομένο ότι:

$$E_q = 0.5 \sum_{n=1}^N (d_{qn} - o_{qn})^2$$

όπου:

$$o_n = \varphi_n \left( \sum_{m=0}^M y_m \right)$$

και

$$y_m = \varphi_m \left( \sum_{l=0}^L w_{lm} x_l \right)$$

Με  $x$  ορίζονται τα διανύσματα εισόδου του νευρωνικού δικτύου, ενώ με  $y$  ορίζεται το διάνυσμα εξόδου της πρώτης στοιβάδας νευρώνων. Τονίζεται εδώ ότι με αρίθμηση 0, ορίζεται το bias για τον συγκεκριμένο νευρώνα. Οπότε, π.χ. για έναν τυχαίο νευρώνα της κρυμμένης στοιβάδας, το bias είναι ίσο με:

$$b_m = w_{0m}x_0$$

όπου ορίζεται ότι  $x_0 = 1$ . Αντίστοιχα ισχύουν και για τους νευρώνες εξόδου.

Με βάση τα παραπάνω προκύπτει ότι:

$$\frac{\partial E_q}{\partial v_{mn}} = -\delta_n y_{qm}$$

όπου  $\delta_n = (d_{qn} - o_{qn})\varphi'_n$  και  $\varphi'_m$  η παράγωγος της συνάρτησης ενεργοποίησης, με βάση το όρισμά της, δηλ. έχοντας

$$u_n = \sum_{m=0}^M y_{qm} v_{mn}$$

η παράγωγος ορίζεται ως

$$\varphi'_n = \frac{d o_{qn}}{d u_n}$$

Για την κρυμμένη στοιβάδα νευρώνων, ισχύουν τα αντίστοιχα, ήτοι:

$$\frac{\partial E_q}{\partial w_{lm}} = -\delta_m y_{ql}$$

όπου όμως ισχύει ότι:

$$\delta_m = \left( \sum_{n=1}^N \delta_n v_{mn} \right) \varphi'_m$$

και όπως και προηγουμένως,  $\varphi'_m$  είναι η παράγωγος της συνάρτησης ενεργοποίησης  $\varphi_m$  ως προς το όρισμά της, δηλ. έχοντας

$$u_m = \sum_{l=0}^L x_l w_{lm}$$

η παράγωγος ορίζεται ως

$$\varphi'_m = \frac{dy_{qm}}{du_m}$$

- **Περίπτωση 2η: Οι τιμές των βαρών και bias διορθώνονται μετά από εισαγωγή όλων των ζευγών διανυσμάτων εισόδου/εξόδου.** Στην προκείμενη περίπτωση, η προσαρμογή των βαρών γίνεται σύμφωνα με τις σχέσεις:

$$\Delta w_{lm} = -a \sum_{q=1}^Q \frac{\partial E_q}{\partial w_{lm}}$$

και

$$\Delta v_{mn} = -a \sum_{q=1}^Q \frac{\partial E_q}{\partial v_{mn}}$$

Ο υπολογισμός των μερικών παραγώγων πραγματοποιείται όπως περιγράφηκε στην πρώτη περίπτωση. Όπως και στην πρώτη περίπτωση, οι νέες τιμές των βαρών του νευρωνικού δικτύου δίνονται από τις παρακάτω σχέσεις:

$$w_{lm}^{new} = w_{lm}^{old} + \Delta w_{lm}$$

$$v_{mn}^{new} = v_{mn}^{old} + \Delta v_{mn}$$

Η διαδικασία εκπαίδευσης συνεχίζεται έως ότου το σφάλμα ελαχιστοποιηθεί. Όταν ολοκληρωθεί, το νευρωνικό δίκτυο πλέον λειτουργεί με τις νέες αποθηκευμένες σταθερές.



Ο αλγόριθμος της οπισθοδρομικής διάδοσης και στις δύο εκδόσεις, του, όπως περιγράφηκε παραπάνω είναι δυνατόν να δώσει ικανοποιητικά αποτελέσματα σε έναν μεγάλο αριθμό προβλημάτων, υπάρχει όμως μία σειρά περιορισμών:

- Σημαντικές παράμετροι όπως ο ρυθμός εκπαίδευσης  $\alpha$  επιλέγονται αυθαίρετα από τον χρήστη, χωρίς να υπάρχει κάποια ικανοποιητική μέθοδος για να γίνει αυτό. Δεδομένης της σημαντικής επίδρασης της παραμέτρου αυτής στα αποτελέσματα που είναι δυνατόν να δώσει ένα νευρωνικό δίκτυο, είναι προφανής η σημασία του περιορισμού αυτού.
- Ακολουθώντας την κατεύθυνση που υποδεικνύει η παράγωγος, με σκοπό την ελαχιστοποίησή της, είναι δυνατόν ο αλγόριθμος να εγκλωβιστεί σε ένα τοπικό ελάχιστο σφάλματος.
- Ο αλγόριθμος οπισθοδρομικής διάδοσης έχει μεγάλο υπολογιστικό κόστος, απαιτώντας μεγάλους χρόνους για την εκπαίδευση του νευρωνικού δικτύου.

Οι παραπάνω περιορισμοί καθιστούν απαραίτητη την επανάληψη της διαδικασίας της εκπαίδευσης πολλές φορές, ώστε να δοκιμαστούν διαφορετικές αρχικές τιμές βαρών και bias. Με τον τρόπο αυτόν, μειώνεται η πιθανότητα εγκλωβισμού σε τοπικό ελάχιστο, καθώς επίσης καθίσταται και δυνατή η δοκιμή εκπαίδευσης για διαφορετικές τιμές του ρυθμού εκπαίδευσης  $\alpha$ .

#### *2.5.2.2 Αλγόριθμος με αριθμητική βελτιστοποίηση*

Οι τεχνικές αριθμητικής βελτιστοποίησης αποσκοπούν στην αύξηση της ταχύτητας σύγκλισης του εκάστοτε αλγορίθμου εκπαίδευσης. Οι συνηθέστερα χρησιμοποιούμενοι αλγόριθμοι αριθμητικής βελτιστοποίησης είναι οι αλγόριθμοι Levenberg-Marquardt και ο αλγόριθμος Quasi-Newton [9].

Για την εφαρμογή της μεθόδου Newton για μη γραμμική βελτιστοποίηση, απαιτείται η χρήση του πίνακα Hessian, ο οποίος περιλαμβάνει τις δεύτερες παραγώγους του μέσου τετραγωνικού σφάλματος του νευρωνικού δικτύου. Ο υπολογισμός όμως των παραγώγων αυτών θα αύξανε σε απαγορευτικό βαθμό το υπολογιστικό κόστος, καθιστώντας την συγκεκριμένη μέθοδο πρακτικά μη εφαρμόσιμη. Απάντηση στον συγκεκριμένο περιορισμό είναι οι αλγόριθμοι Quasi-Newton, η περιγραφή των οποίων ξεφεύγουν από τα πλαίσια της παρούσας εργασίας.

Ο αλγόριθμος Levenberg-Marquardt, ο οποίος χρησιμοποιείται κατά κύριο λόγο στην παρούσα εργασία, είναι σχεδιασμένος για να έχει ταχύτητα σύγκλισης αντίστοιχη με την ταχύτητα σύγκλισης της μεθόδου Newton, χωρίς όμως να απαιτείται αναλυτικός υπολογισμός του πίνακα Hessian.

Στον αλγόριθμο Levenberg-Marquardt, ο πίνακας Hessian, αντικαθίσταται από την παρακάτω προσέγγιση:

$$H = J^T J$$

Η παράγωγος του μέσου τετραγωνικού σφάλματος εκφράζεται από την παρακάτω σχέση:

$$G = J^T E$$

Όπου:

$G$ : Η κλίση (κατεύθυνση μέγιστης μείωσης σφάλματος) του μέσου τετραγωνικού σφάλματος σε σχέση με τα βάρη και τα bias.

$J$ : Ο ιακωβιανός πίνακας που περιέχει τις πρώτες παραγώγους του μέσου τετραγωνικού σφάλματος του νευρωνικού δικτύου, σε σχέση με τα βάρη και τα bias.

$E$ : Το διάνυσμα με τις τιμές του μέσου τετραγωνικού σφάλματος του νευρωνικού δικτύου.

Κατά την εκτέλεση του αλγορίθμου Levenberg-Marquardt, η μεταβολή των βαρών και των bias δίνεται από την σχέση:

$$\Delta w = -[H + \mu I]^{-1} G$$

Με βάση την προσέγγιση που αναφέρθηκε παραπάνω για τον πίνακα Hessian, προκύπτει ότι:

$$\Delta w = -[J^T J + \mu I]^{-1} J^T E$$

όπου ο συντελεστής  $\mu$  καθορίζει την ταχύτητα σύγκλισης του αλγορίθμου, και μεταβάλλεται κατά την διάρκεια της εκπαίδευσης. Συντελεστής  $\mu = 0$  σημαίνει ότι χρησιμοποιείται η μέθοδος Newton με χρήση προσέγγισης για τον πίνακα Hessian. Μεγάλες τιμές του συντελεστή έχουν σαν αποτέλεσμα μία αντίστοιχη μέθοδο εκπαίδευσης με οπισθοδρομική διάδοση. Καθώς, είναι επιθυμητή η μετάβαση στην μέθοδο Newton λόγω της ταχύτητας σύγκλισής της, ο συντελεστής  $\mu$  μειώνεται προς το μηδέν κατά την διάρκεια της εκπαίδευσης.

Παρ' όλο που η χρήση του αλγορίθμου εκπαίδευσης Levenberg-Marquardt έχει το μειονέκτημα ότι στις περιπτώσεις μεγάλων συστημάτων καταλήγει σε μεγάλους πίνακες, οπότε και χρησιμοποιεί αυξημένη ποσότητα μνήμης του υπολογιστή, είναι από τους συνηθέστερα χρησιμοποιούμενους.

## **2.6 Ικανότητα γενίκευσης αποτελεσμάτων νευρωνικών δικτύων**

Όπως αναφέρθηκε στις προηγούμενες παραγράφους, κατά την διάρκεια της επιτηρούμενης εκπαίδευσης, τα νευρωνικά δίκτυα εκπαιδεύονται πάνω σε μία ομάδα από ζεύγη δεδομένων εισόδου/εξόδου για την εκάστοτε κατεργασία. Η πρακτική χρησιμότητα των νευρωνικών δικτύων εξαρτάται από τον βαθμό στον οποίο είναι σε θέση να πραγματοποιούν προβλέψεις για νέα δεδομένα, τα οποία δεν χρησιμοποιήθηκαν κατά την διαδικασία της εκπαίδευσης. Η ικανότητα αυτή των νευρωνικών δικτύων, ονομάζεται ικανότητα γενίκευσης [8], [9].

Η ικανότητα γενίκευσης των αποτελεσμάτων που παράγει ένα νευρωνικό δίκτυο εξαρτάται από την πολυπλοκότητα του προβλήματος, τον αριθμό των διαθέσιμων δεδομένων εκπαίδευσης, καθώς επίσης και τον αριθμό των παραμέτρων του δικτύου, ο οποίος εξαρτάται άμεσα από τον αριθμό νευρώνων του δικτύου. Π.χ. μικρός αριθμός δεδομένων, συνδυασμένος με δίκτυο μεγάλης πολυπλοκότητας, θα οδηγήσει σε ένα συχνό πρόβλημα που εμφανίζεται κατά την εκπαίδευση των νευρωνικών δικτύων και ονομάζεται *overfitting*. Όταν παρουσιάζεται το συγκεκριμένο πρόβλημα, πρακτικά σημαίνει ότι το δίκτυο απομνημόνευσε τα παραδείγματα εκπαίδευσης και δεν δύναται να εκτελέσει με αποδεκτή ακρίβεια γενίκευση αποτελεσμάτων.

Μία από τις συνηθέστερες μεθόδους της αποτροπής του overfitting, η οποία και χρησιμοποιείται στην παρούσα εργασία, είναι η μέθοδος της πρόωρης διακοπής. Κατά την μέθοδο πρόωρης διακοπής, τα δεδομένα χωρίζονται σε τρία υποσύνολα, αποτελούμενα από ζεύγη δεδομένων εισόδου/εξόδου

- Το πρώτο υποσύνολο χρησιμοποιείται για τον υπολογισμό των βαρών και των bias του νευρωνικού δικτύου, με μία από τις μεθόδους που αναφέρθηκαν στις προηγούμενες παραγράφους.
- Το δεύτερο υποσύνολο χρησιμοποιείται για την επικύρωση (validation) των αποτελεσμάτων της εκπαίδευσης. Δεν χρησιμοποιείται για τον υπολογισμό των βαρών και των bias, αλλά το σφάλμα πρόβλεψης του δικτύου σε σχέση με τα δεδομένα αυτά, παρακολουθείται κατά την διάρκεια της εκπαίδευσης. Εφόσον κατά την διάρκεια της εκπαίδευσης μειώνεται, η ικανότητα γενίκευσης του δικτύου αυξάνεται. Εφόσον το σφάλμα αρχίσει να αυξάνεται, αυτό είναι ένδειξη ότι το δίκτυο χάνει την ικανότητα γενίκευσης, εισάγοντας τον κίνδυνο του overfitting, οπότε ο αλγόριθμος εκπαίδευσης διακόπτεται.
- Η τρίτη ομάδα είναι η ομάδα ελέγχου και όπως και η δεύτερη επίσης δεν χρησιμοποιείται για τον υπολογισμό των βαρών και των bias του νευρωνικού δικτύου. Είναι μία ξεχωριστή ομάδα δεδομένων η οποία χρησιμοποιείται κατά κύριο λόγο για την σύγκριση διαφορετικών μεθόδων και μοντέλων.

Τονίζεται εδώ ότι οι μέθοδος για την αποτροπή του overfitting που περιγράφεται παραπάνω, είναι δυνατόν να εφαρμοστεί σε οποιονδήποτε αλγόριθμο εκπαίδευσης.

### 3. Νευρωνικά δίκτυα ακτινικής βάσης (radial basis function neural networks)

Πρόκειται για μία ειδική περίπτωση νευρωνικών δικτύων των οποίων η αρχιτεκτονική διαφέρει από την συνηθισμένη και, ως εκ τούτου, έχουν διαφορετικό τρόπο λειτουργίας. Τα νευρωνικά δίκτυα ακτινικής βάσης παρουσιάζουν μία σειρά πλεονεκτημάτων που συνίστανται στη μεγαλύτερη ευκολία κατασκευής τους, χωρίς να μειώνεται η ακρίβεια επίλυσης του αντιμετωπιζόμενου προβλήματος. Τα πλεονεκτήματα αυτά ήταν αρκετά για να στραφεί η έρευνα προς την κατεύθυνση αυτή και να σημειωθεί σημαντική πρόοδος στις συγκεκριμένες τεχνικές [11].

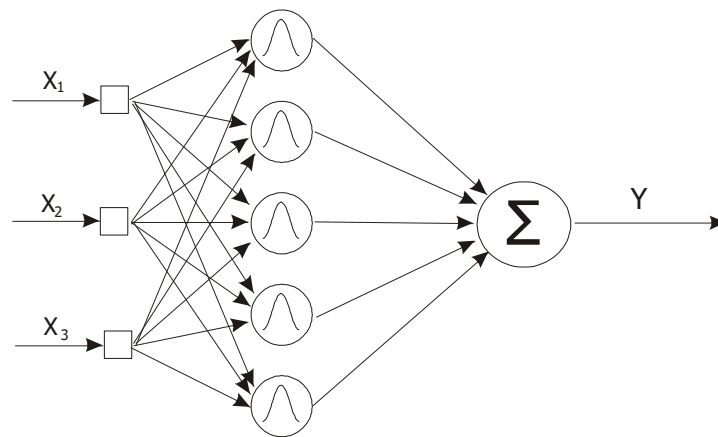
Οι εφαρμογές των νευρωνικών δικτύων ακτινικής βάσης είναι πολλές, το συνηθέστερο όμως πεδίο εφαρμογής τους σχετίζεται με την μοντελοποίηση κάποιου συστήματος. Παρακάτω αναφέρονται ενδεικτικά δύο εφαρμογές, η πρώτη αφορά την μαθηματική μοντελοποίηση του ποσοστού της ανεργίας σε συγκεκριμένη χώρα ενώ στη δεύτερη κατασκευάζεται νευρωνικό δίκτυο το οποίο κρίνει αν μία βιομηχανική κατεργασία διεξάγεται ομαλά και αν όχι εντοπίζει το είδος της δυσλειτουργίας.

1. *Μαθηματική μοντελοποίηση της τιμής του ποσοστού της ανεργίας στην δυτική Γερμανία την περίοδο 1948-1980*: Η συγκεκριμένη εργασία δημοσιεύτηκε το 1991 από τους συγγραφείς S. Chen, C. F. N. Cowan, P. M. Grant [12]. Μία από τις εφαρμογές που παρουσιάστηκαν αφορούσε την κατασκευή νευρωνικού δικτύου ακτινικής βάσης το οποίο καθιστούσε δυνατή την πρόβλεψη της τιμής του ποσοστού της ανεργίας για ένα μήνα στην δυτική Γερμανία με βάση τις τιμές του τους προηγούμενους 13 μήνες. Κατασκευάστηκε δίκτυο με 55 νευρώνες οι οποίοι χρησιμοποιούσαν την συνάρτηση thin plate spline. Το δίκτυο είχε προφανώς  $L=13$  εισόδους και  $M=1$  έξοδο.
2. *Εντοπισμός δυσλειτουργιών κατεργασίας (failure detection)*: Μία τέτοια εργασία δημοσιεύτηκε το 1990 από τους ερευνητές James A. Leonard και Mark A. Kramer [11]. Παρουσιάστηκε η χρήση των νευρωνικών δικτύων ακτινικής βάσης για τον σκοπό αυτό, η οποία κατέληξε στην ανάπτυξη δικτύου του οποίου η έξοδος αποτελείτο από ένα διάνυσμα, κάθε διάσταση του οποίου αντιστοιχούσε σε ένα είδος δυσλειτουργίας. Μία τελευταία διάσταση αντιστοιχούσε στην ομαλή λειτουργία. Ανάλογα με το είδος

λειτουργίας, το διάνυσμα εξόδου είχε τον αριθμό 1 στην αντίστοιχη διάσταση και 0 στις υπόλοιπες. Η συνάρτηση που χρησιμοποιήθηκε ήταν η συνάρτηση Gauss.

### 3.1 Δομή και τρόπος λειτουργίας των νευρωνικών δικτύων ακτινικής βάσης

Αποτελούνται από τρεις στοιβάδες νευρώνων, καθεμία εκ των οποίων επιτελεί διαφορετικό ρόλο, και τις μεταξύ τους συνδέσεις. Ένα νευρωνικό δίκτυο ακτινικής βάσης με τρεις εισόδους, πέντε νευρώνες στην πρώτη στοιβάδα και μία έξοδο εικονίζεται στο σχήμα 29.



Σχήμα 29: Νευρωνικό δίκτυο ακτινικής βάσης με τρεις εισόδους πέντε νευρώνες στην κρυμμένη στοιβάδα και μία έξοδο

Ο τρόπος της λειτουργίας των νευρωνικών δικτύων ακτινικής βάσης θα αναλυθεί μέσω της παρουσίασης του τρόπου χρήσης τους κατά την προσέγγιση πειραματικών δεδομένων πολυμεταβλητών συστημάτων [11].

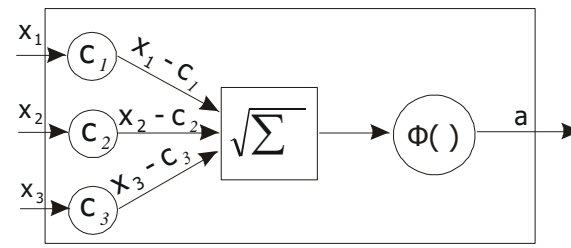
Αν υποθεθεί ότι πρέπει να προσεγγιστούν  $K$  πειραματικές μετρήσεις με  $L$  μεταβλητές εισόδου (input variables) οι οποίες δίνονται με την μορφή των διανύσματος  $\mathbf{x}(k)=[x_1(k) \ x_2(k) \ \dots \ x_l(k) \ \dots \ x_L(k)]$  και  $M$  μεταβλητές εξόδου (output) οι οποίες δίνονται με την μορφή του διανύσματος  $\mathbf{y}(k)=[y_1(k) \ y_2(k) \ \dots \ y_m(k) \ \dots \ y_M(k)]$ , όπου  $k=1,2,\dots,K$ ,  $l=1,2,\dots,L$  και  $m=1,2,\dots,M$  τότε προκύπτουν  $K$  διανύσματα εισόδου και εξόδου από τα οποία προκύπτει ο πίνακας των τιμών εισόδου  $\mathbf{X}$  και ο αντίστοιχος πίνακας των τιμών εξόδου  $\mathbf{Y}$ .

$$\mathbf{X} = \begin{bmatrix} x_1(1) & x_2(1) & \cdots & x_L(1) \\ x_1(2) & x_2(2) & \cdots & x_L(2) \\ \vdots & \vdots & \ddots & \vdots \\ x_1(\mathbf{K}) & x_2(\mathbf{K}) & \cdots & x_L(\mathbf{K}) \end{bmatrix}, \mathbf{Y} = \begin{bmatrix} y_1(1) & y_2(1) & \cdots & y_M(1) \\ y_1(2) & y_2(2) & \cdots & y_M(2) \\ \vdots & \vdots & \ddots & \vdots \\ y_1(\mathbf{K}) & y_2(\mathbf{K}) & \cdots & y_M(\mathbf{K}) \end{bmatrix}$$

Ένα νευρωνικό δίκτυο ακτινικής βάσης το οποίο προσεγγίζει τα πειραματικά δεδομένα εξόδου  $\mathbf{Y}$  με χρήση των δεδομένων εισόδου  $\mathbf{X}$  αποτελείται από τα εξής δομικά στοιχεία [11]:

- **Στοιβάδα εισόδου (input layer):** Πρόκειται για το τμήμα του δικτύου το οποίο δέχεται τα δεδομένα εισόδου. Αποτελείται από  $L$  νευρώνες καθένας εκ των οποίων δέχεται από μία μεταβλητή εισόδου. Έτσι π.χ. για την δεύτερη σειρά δεδομένων ο πρώτος νευρώνας θα δεχθεί την μεταβλητή  $x_1(2)$ , ο δεύτερος την μεταβλητή  $x_2(2)$  κ.ο.κ.
- **Συνδέσεις (connections) μεταξύ πρώτης και δεύτερης στοιβάδας νευρώνων:** Οι συνδέσεις αυτές μεταφέρουν τα δεδομένα από την πρώτη στοιβάδα νευρώνων στη δεύτερη. Κάθε νευρώνας της πρώτης στοιβάδας συνδέεται, οπότε και μεταφέρει τα δεδομένα του, σε όλους τους νευρώνες της δεύτερης στοιβάδας. Με τον τρόπο αυτόν κάθε νευρώνας της δεύτερης στοιβάδας δέχεται δεδομένα από όλους τους νευρώνες της πρώτης. Για παράδειγμα, ο δεύτερος νευρώνας του πρώτου στρώματος θα μεταφέρει το δεδομένο  $x_2(k)$  που αντιστοιχεί στην  $k$  πειραματική μέτρηση σε όλους τους νευρώνες της δεύτερης. Αντίστοιχα όλοι οι νευρώνες της δεύτερης στοιβάδας θα δεχθούν το διάνυσμα  $\mathbf{x}(k)$ .
- **Κρυφή στοιβάδα νευρώνων (hidden layer):** Η στοιβάδα αυτή δεν συμμετέχει στην εισαγωγή ή στην εξαγωγή δεδομένων από το δίκτυο αλλά εκτελεί εσωτερικές πράξεις του, και πρόκειται για το κύριο δομικό στοιχείο του. Αποτελείται από  $N$  κρυμμένους νευρώνες (hidden nodes) καθένας εκ των οποίων δέχεται σαν είσοδο, όπως αναφέρθηκε παραπάνω, το διάνυσμα  $\mathbf{x}_k$ . Ο αριθμός τους,  $N$ , δεν καθορίζεται από τα δεδομένα αλλά προσδιορίζεται με διάφορες μεθόδους, όπως θα αναλυθεί παρακάτω. Η έξοδος κάθε νευρώνα της

στοιβάδας αυτής προκύπτει από την διαδικασία που εικονίζεται παραστατικά στο σχήμα 2.5. Αρχικά αφαιρείται από το διάνυσμα εισόδου  $\mathbf{x}$  ένα διάνυσμα  $\mathbf{c}_n$  το οποίο είναι χαρακτηριστικό για τον νευρώνα  $n$ . Το διάνυσμα  $\mathbf{c}_n = [c_{n,1} \ c_{n,2} \ \dots \ c_{n,l} \ \dots \ c_{n,L}]$  διάστασης  $L$  είναι το κέντρο του νευρώνα  $n$ , αποτελεί το κυριότερο χαρακτηριστικό του και υπολογίζεται με διάφορους τρόπους, όπως θα δειχθεί παρακάτω. Τα προκύπτοντα στοιχεία του διανύσματος  $\mathbf{x} - \mathbf{c}_n$  υψώνονται στο τετράγωνο, αθροίζονται και υπολογίζεται η τετραγωνική ρίζα του αθροίσματος (δηλ. υπολογίζεται η ευκλείδεια νόρμα του  $\mathbf{x} - \mathbf{c}_n$ ). Στην συνέχεια εφαρμόζεται μία μη γραμμική σχέση που επιλέγεται ανάλογα με το πρόβλημα. Στο σχήμα 30 το διάνυσμα  $\mathbf{x}$  αποτελείται μόνον από τρία στοιχεία, τα  $x_1$ ,  $x_2$  και  $x_3$  και το διάνυσμα  $\mathbf{c}_n$  από τα  $c_1$ ,  $c_2$  και  $c_3$ .



Σχήμα 30: Νευρώνας κρυμμένης στοιβάδας νευρωνικού δικτύου ακτινικής βάσης

Το σύνολο των κέντρων όλων των νευρώνων του δικτύου απεικονίζεται στον πίνακα  $\mathbf{C}$ , κάθε γραμμή του οποίου αντιστοιχεί και σε ένα κέντρο  $\mathbf{c}_n$ .

$$\mathbf{C} = \begin{bmatrix} c_{1,1} & c_{1,2} & \dots & c_{1,L} \\ c_{2,1} & c_{2,2} & \dots & c_{2,L} \\ \vdots & \vdots & \ddots & \vdots \\ c_{N,1} & c_{N,2} & \dots & c_{N,L} \end{bmatrix}$$

Οι πιο συχνά χρησιμοποιούμενες μη γραμμικές συναρτήσεις είναι η Gauss και η thin plate spline [11], [12], [13].

- *Συνάρτηση Gauss*: Έχει την παρακάτω μορφή και είναι παρόμοια με την κατανομή Gauss από την οποία άλλωστε έχει ληφθεί και η ονομασία.

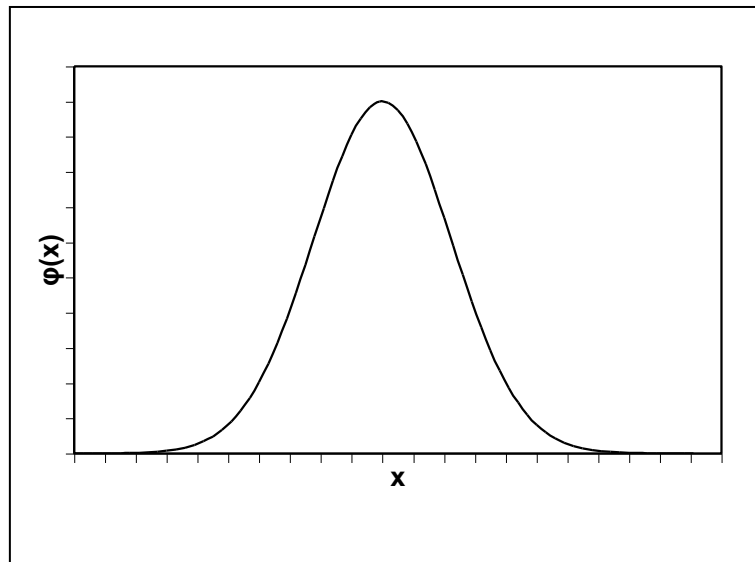
$$\phi(x) = \exp\left(-\frac{x^2}{\sigma^2}\right)$$



Η παραπάνω σχέση, όπως άλλωστε και η κατανομή Gauss έχει γραφική παράσταση της μορφής που παρουσιάζεται στο σχήμα 31. Όταν χρησιμοποιείται, η έξοδος του νευρώνα  $n$  για το  $k$  δεδομένο εισόδου δίνεται από την παρακάτω σχέση.

$$a_n(k) = \exp\left(\frac{-\|\mathbf{x}(k) - \mathbf{c}_n\|^2}{\sigma_n^2}\right)$$

Με  $a_n(k)$  συμβολίζεται η έξοδος του  $n$ -οστού νευρώνα, όταν του δοθεί σαν είσοδος το πειραματικό δεδομένο  $\mathbf{x}(k)$ , με  $n=1, 2, \dots, N$ .



**Σχήμα 31: : Γραφική παράσταση της συνάρτησης Gauss**

Η σταθερά  $\sigma_n$  εμφανίζεται μόνο όταν χρησιμοποιείται η σχέση Gauss, και υπολογίζεται συνήθως από την παρακάτω σχέση

$$\sigma_n = \frac{1}{p} \cdot \sum_{h=1}^p \|\mathbf{c}_n - \mathbf{c}_h\|^2$$

όπου  $\mathbf{c}_h$ ,  $h=1,2,\dots,p$  τα  $p$  κοντινότερα κέντρα στο  $\mathbf{c}_n$  με βάση την ευκλείδεια νόρμα. Ο αριθμός  $p$  είναι πάντοτε θετικός ακέραιος και επιλέγεται από τον χρήστη κατά την διαδικασία εκπαίδευσης του δικτύου.

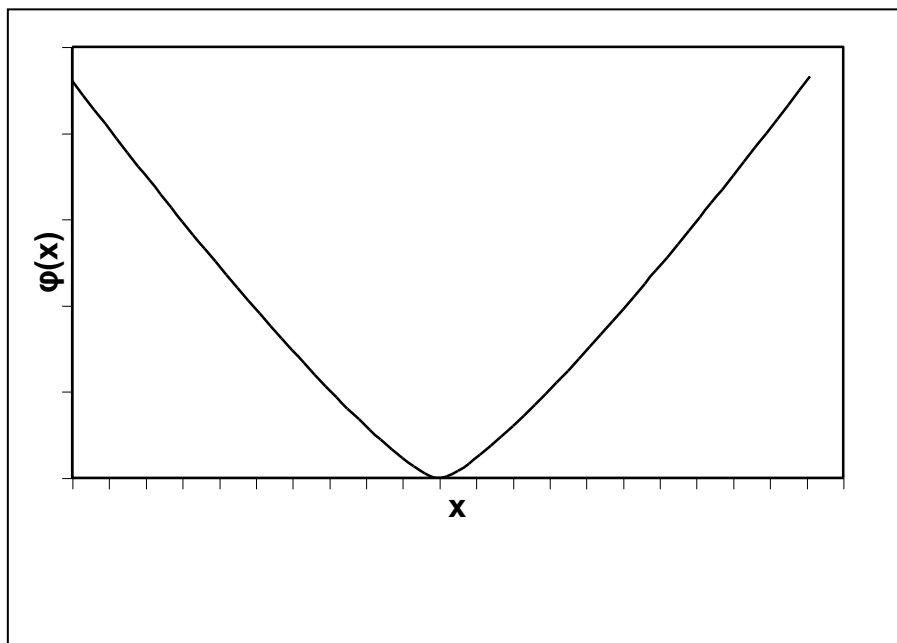
- Σχέση *thin plate spline*: Μία επίσης πολύ συχνά χρησιμοποιούμενη σχέση είναι η *thin plate spline*:

$$\phi(x) = x^2 \cdot \log(x)$$

Η εν λόγω σχέση παρουσιάζεται διαγραμματικά στο σχήμα 32 και η έξοδος του νευρώνα δίδεται από την παρακάτω σχέση:

$$a_n(k) = v_n^2(k) \cdot \log(v_n(k))$$

όπου  $v_n(k) = \|\mathbf{x}(k) - \mathbf{c}_n\|$ . Η σχέση αυτή πλεονεκτεί σε ευκολία χρήσης, διότι η μοναδική παράμετρος που πρέπει να προσδιορισθεί είναι ο πίνακας  $\mathbf{C}$ , χωρίς να μειονεκτεί σε ακρίβεια προσέγγισης του προβλήματος.



Σχήμα 32: Γραφική παράσταση της συνάρτησης *thin plate spline*

Οι έξοδοι της δεύτερης στοιβάδας των νευρώνων απεικονίζονται στον  $K \times N$  πίνακα  $\mathbf{A}$ :

$$\mathbf{A} = \begin{bmatrix} a_1(1) & a_2(1) & \cdots & a_N(1) \\ a_1(2) & a_2(2) & \cdots & a_N(2) \\ \vdots & \vdots & \ddots & \vdots \\ a_1(K) & a_2(K) & \cdots & a_N(K) \end{bmatrix}$$

- **Συνδέσεις μεταξύ δεύτερης και τρίτης στοιβάδας νευρώνων:** Μεταφέρουν τις εξόδους από τους νευρώνες της δεύτερης στοιβάδας στους νευρώνες της τρίτης, αφού τους πολλαπλασιάσουν με έναν αριθμό  $w_{n,m}$  που ονομάζεται βάρος. Με  $w_{n,m}$  συμβολίζεται το βάρος με το οποίο πολλαπλασιάζεται η έξοδος του  $n$ -οστού νευρώνα της δεύτερης στοιβάδας που μεταφέρεται μέσω της σύνδεσης στον  $m$ -οστό νευρώνα της τρίτης στοιβάδας. Κάθε νευρώνας της δεύτερης στοιβάδας συνδέεται, οπότε και μεταφέρει τα δεδομένα του σε όλους τους νευρώνες της τρίτης. Αντίστροφα, κάθε νευρώνας της τρίτης δέχεται δεδομένα από όλους τους νευρώνες της δεύτερης. Τα βάρη όλων των συνδέσεων συνιστούν τον  $N \times M$  πίνακα  $\mathbf{W}$ .

$$\mathbf{W} = \begin{bmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,M} \\ w_{2,1} & w_{2,2} & \cdots & w_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ w_{N,1} & w_{N,2} & \cdots & w_{N,M} \end{bmatrix}$$

- **Στοιβάδα εξόδου (output layer):** Κάθε νευρώνας της τρίτης στοιβάδας δέχεται όλες τις εξόδους των νευρώνων του δεύτερου πολλαπλασιασμένες επί τα βάρη  $w_{n,m}$ , και τις αθροίζει. Κατά συνέπεια η τελική έξοδος  $\mathbf{Y}_{\text{net}}$  του νευρωνικού δικτύου είναι το γινόμενο του πίνακα  $\mathbf{A}$  επί τον  $\mathbf{W}$ :

$$\mathbf{Y}_{\text{net}} = \mathbf{A} \cdot \mathbf{W}$$

Η ικανότητα του δικτύου να προσεγγίζει δεδομένα προσδιορίζεται συνήθως από το άθροισμα των τετραγώνων των σφαλμάτων,  $SSE$  που ορίζεται από την παρακάτω σχέση:

$$SSE = \|\mathbf{Y}_{net} - \mathbf{Y}\|^2$$

Το μέγεθος αυτό αντιπροσωπεύει την ευκλείδεια νόρμα της διαφοράς (σφάλμα) μεταξύ των τιμών των μεταβλητών που μετρήθηκαν πειραματικά και των τιμών των μεταβλητών που βρέθηκαν από το δίκτυο.

Ένας άλλος τρόπος προσδιορισμού της ικανότητας του δικτύου να προσεγγίζει τα δεδομένα του εξεταζόμενου προβλήματος είναι το μέσο σφάλμα,  $ME$ , το οποίο υπολογίζεται από την παρακάτω σχέση:

$$ME = mean(abs(\mathbf{Y} - \mathbf{Y}_{net}))$$

Το μέσο σφάλμα εκφράζει την μέση τιμή της απόλυτης τιμής της διαφοράς μεταξύ των πειραματικών τιμών εξόδου και των τιμών εξόδου που προέκυψαν από το δίκτυο. Όπως και στο  $SSE$ , εάν η έξοδος είναι μίας μεταβλητής (δηλαδή οι πίνακες  $\mathbf{Y}$  και  $\mathbf{Y}_{net}$  είναι πίνακες στήλες) τότε το  $ME$  είναι αριθμός, ενώ εάν είναι  $M$  μεταβλητών το  $ME$  θα είναι πίνακας γραμμή διάστασης  $M$  κάθε στήλη του οποίου αντιπροσωπεύει το μέσο σφάλμα της αντίστοιχης μεταβλητής. Συχνά είναι χρήσιμο να ορίζεται το μέσο σφάλμα ως ποσοστό επί της τιμής του μετρούμενου μεγέθους.

Πρέπει εδώ να παρατηρηθεί ότι τα δεδομένα σταθμίζονται (scaling) κατάλληλα προτού εισαχθούν στο νευρωνικό δίκτυο ώστε να είναι της ίδιας τάξης μεγέθους. Αν για παράδειγμα μοντελοποιείται ένας χημικός αντιδραστήρας όπου υπεισέρχονται συγκεντρώσεις της τάξης 0.1-1M και θερμοκρασίες της τάξης 300-400K τα δεδομένα σταθμίζονται κατάλληλα ώστε και τα δύο μεγέθη να παίρνουν τιμές από 0 έως π.χ. 1, 10, 100. Αφού εξαχθούν τα αποτελέσματα από το δίκτυο πραγματοποιείται η αντίστροφη διαδικασία ώστε να ληφθούν οι πραγματικές τιμές. Στην διαδικασία της στάθμισης θα γίνει εκτενέστερη αναφορά παρακάτω.

### 3.2 Εκπαίδευση νευρωνικών δικτύων ακτινικής βάσης

Τα νευρωνικά δίκτυα ακτινικής βάσης, όπως και τα νευρωνικά δίκτυα γενικότερα ανήκουν στις μεθόδους τεχνητής νοημοσύνης. Κατά συνέπεια εκπαιδεύονται στην συμπεριφορά του προς εξέταση συστήματος υπό ορισμένες συνθήκες για να είναι σε θέση να προσδιορίσουν τον τρόπο με τον οποίο θα συμπεριφερθεί υπό άλλες συνθήκες.

Με τον όρο εκπαίδευση νευρωνικού δικτύου ακτινικής βάσης εννοείται ο προσδιορισμός των παραμέτρων του, ήτοι, του αριθμού των νευρώνων της δεύτερης στοιβάδας  $N$ , των κέντρων  $c_n$ , των βαρών  $w_{n,m}$  με  $n=1,2,\dots,N$  και  $m=1,2,\dots,M$  και των σταθερών  $\sigma_n$  (εάν χρησιμοποιείται η συνάρτηση Gauss). Αυτό γίνεται με προσαρμογή των μεγεθών αυτών στην συμπεριφορά του προς εξέταση συστήματος, η οποία ορίζεται από μια ομάδα τιμών εισόδου  $\mathbf{X}$  και εξόδου  $\mathbf{Y}$  που προσδιορίστηκαν πειραματικά. Η προσαρμογή αυτή περιλαμβάνει την εφαρμογή κάποιας διαδικασίας που περιλαμβάνει εκτέλεση ενός μαθηματικού αλγορίθμου.

- **Αριθμός νευρώνων κρυφής στοιβάδας,  $N$ :** Όσο μεγαλύτερος είναι, τόσο αυξάνεται η δυνατότητα του δικτύου να προσεγγίζει τις διακυμάνσεις των μεταβλητών εξόδου. Εάν όμως είναι υπερβολικά μεγάλος, η έξοδος του δικτύου θα ακολουθεί ανεπιθύμητες διακυμάνσεις όπως θόρυβο και πειραματικά σφάλματα οπότε η δυνατότητα πρόβλεψης νέων μεταβλητών εξόδου θα είναι μειωμένη. Ο βέλτιστος αριθμός των νευρώνων ποικίλλει από πρόβλημα σε πρόβλημα και είναι η σημαντικότερη παράμετρος κατά την κατασκευή ενός μαθηματικού μοντέλου με νευρωνικά δίκτυα ακτινικής βάσης. Ο συνηθέστερος τρόπος προσδιορισμού του είναι με δοκιμή και σφάλμα, όπου σημαντικό ρόλο παίζει η εμπειρία από την επίλυση άλλων αντίστοιχων προβλημάτων. Κατά καιρούς όμως έχουν δημοσιευτεί εργασίες όπου προτείνονται συγκεκριμένες μέθοδοι υπολογισμού του αριθμού αυτού αυξάνοντας τον βαθμό αυτοματοποίησης κατά την κατασκευή ενός μαθηματικού μοντέλου με το συγκεκριμένο είδος νευρωνικών δικτύων.
- **Κέντρα των νευρώνων  $c_n$ :** Πρόκειται για τον σημαντικότερο παράγοντα μετά την επιλογή του αριθμού των κέντρων  $N$  όσον αφορά την ικανότητα ενός δικτύου να προσεγγίζει ένα πρόβλημα και έχουν δημοσιευτεί πολλές εργασίες για τον σκοπό αυτό. Ο απλούστερος –άρα και λιγότερο ακριβής, είναι η επιλογή

των κέντρων μέσα από τα πειραματικά δεδομένα εισόδου. Τις περισσότερες φορές όμως χρησιμοποιείται ένας μαθηματικός αλγόριθμος ένας εκ των οποίων παρουσιάζεται ακολούθως. Η συγκεκριμένη διαδικασία ονομάζεται k-means clustering και έχει δημοσιευτεί σε αρκετές εργασίες με πολλές παραλλαγές. Η διαδικασία που ακολουθείται κατά την εφαρμογή του είναι η εξής [11]:

1. Έχοντας επιλέξει τον αριθμό των νευρώνων του δεύτερου στρώματος,  $N$ , επιλέγονται τυχαία  $N$  δεδομένα εισόδου από τα  $K$  διανύσματα  $\mathbf{x}_k$  και ορίζονται σαν οι αρχικές τιμές των κέντρων  $\mathbf{c}_n^0$  του νευρωνικού δικτύου με  $n=1,2,\dots,N$ .

Τα δύο επόμενα βήματα επαναλαμβάνονται  $G$  φορές, όπου  $G$  ένας προεπιλεγμένος ακέραιος αριθμός.

2. Για κάθε δεδομένο  $\mathbf{x}_k$ ,  $k=1,2,\dots,K$  βρίσκεται το κοντινότερό του κέντρο  $\mathbf{c}_n$ ,  $n=1,2,\dots,N$  με βάση την ευκλείδεια νόρμα  $\|\mathbf{x}_k - \mathbf{c}_n^{(g-1)}\|$  όπου  $g=1,2,\dots,G$ , ο αριθμός της επανάληψης.
3. Κάθε κέντρο  $\mathbf{c}_n$  μετατοπίζεται προς το δεδομένο  $\mathbf{x}_k$ , στο οποίο βρέθηκε κοντινότερο. Η μετατόπιση γίνεται κατά έναν συντελεστή  $\alpha$ , σύμφωνα με τη σχέση:

$$\mathbf{c}_n^g = \mathbf{c}_n^{g-1} + \alpha \cdot (\mathbf{x}_k - \mathbf{c}_n^{(g-1)})$$

Εάν χρησιμοποιείται η σχέση Gauss υπολογίζονται στην συνέχεια οι σταθερές  $\sigma_n$ . Ο αριθμός  $p$  επιλέγεται πάντοτε θετικός ακέραιος της τάξης του 2 με 3

- **Υπολογισμός βαρών,  $\mathbf{W}$ :** Μετά την επανάληψη των δύο παραπάνω βημάτων  $G$  φορές, ακολουθεί ο υπολογισμός του πίνακα των βαρών  $\mathbf{W}$ . Τα βάρη υπολογίζονται έτσι ώστε να ελαχιστοποιείται η ευκλείδεια νόρμα της διαφοράς μεταξύ των πειραματικών δεδομένων εξόδου  $\mathbf{Y}$  και των τιμών εξόδου  $\mathbf{Y}_{\text{net}}$  που προκύπτουν από το δίκτυο. Αποδεικνύεται από την θεωρία των ελαχίστων τετραγώνων, ότι ο πίνακας  $\mathbf{W}$  με την ιδιότητα αυτή δίνεται από την παρακάτω σχέση:

$$\mathbf{W} = (\mathbf{A}^T \cdot \mathbf{A})^{-1} \cdot (\mathbf{A}^T \cdot \mathbf{Y})$$

Συνήθως δεν χρησιμοποιούνται όλα τα πειραματικά δεδομένα εισόδου και εξόδου του συστήματος για την εκπαίδευση του δικτύου, αλλά μόνο ένα μέρος τους. Τα υπόλοιπα χρησιμοποιούνται για να ελεγχθεί αν το δίκτυο παράγει σωστά αποτελέσματα.

Τα δεδομένα εισόδου, όπως ειπώθηκε παραπάνω, προτού εισαχθούν στο δίκτυο σταθμίζονται ώστε να έχουν την ίδια τάξη μεγέθους και να κινούνται οι τιμές τους μέσα σε πλαίσια τα οποία επιλέγονται από τον χρήστη του δικτύου. Ο απλούστερος τρόπος στάθμισης περιλαμβάνει γραμμική παρεμβολή δηλαδή πραγματοποιείται με χρήση της παρακάτω:

$$X_l^{scaled}(k) = X_l^{min,scaled}(k) + (X_l^{real}(k) - X_l^{min,real}(k)) \cdot \frac{X_l^{max,scaled}(k) - X_l^{min,scaled}(k)}{X_l^{max,real}(k) - X_l^{min,real}(k)}$$

Ο εκθέτης *scaled* δηλώνει την σταθμισμένη τιμή του μεγέθους  $x_l(k)$  ενώ ο εκθέτης *real* δηλώνει την πραγματική του τιμή. Ο εκθέτης *min* στις πραγματικές τιμές δηλώνει την ελάχιστη τιμή που παρουσιάζεται στα πειραματικά δεδομένα ενώ αντίθετα ο εκθέτης *max* την μέγιστη. Όσον αφορά τις σταθμισμένες τιμές, η μέγιστη και η ελάχιστη τιμή ορίζονται κατά βούληση ώστε να έχουμε την επιθυμητή τάξη μεγέθους.

Αφού χρησιμοποιηθεί το δίκτυο, θα παραχθούν τιμές εξόδου  $y_m(k)$  οι οποίες θα είναι σταθμισμένες, οπότε οι πραγματικές τιμές θα δοθούν από την παρακάτω σχέσης η οποία αποτελεί αντίστροφη της τελευταίας σχέσης σχέσης.

$$Y_m^{real}(k) = Y_m^{min,real}(k) + (Y_m^{scaled}(k) - Y_m^{min,scaled}(k)) \cdot \frac{Y_m^{max,real}(k) - Y_m^{min,real}(k)}{Y_m^{max,scaled}(k) - Y_m^{min,scaled}(k)}$$

Οι εκθέτες  $min$  και  $max$  στις σταθμισμένες τιμές του μεγέθους  $y_m(k)$  αντιπροσωπεύουν την μέγιστη και την ελάχιστη τιμή που έχει ορισθεί από τον χρήστη (συνήθως ίδιες με τις μέγιστες και ελάχιστες τιμές που ορίζονται για την είσοδο). Ως μέγιστη και ελάχιστη τιμή στις πραγματικές τιμές όμως λαμβάνονται η μέγιστη και η ελάχιστη τιμή που λαμβάνει το μέγεθος αυτό στα πειραματικά δεδομένα εκπαίδευσης του δικτύου.



## 4. Μαθηματική μοντελοποίηση λείανσης μετάλλων με χρήση νευρωνικών δικτύων

Στο παρόν κεφάλαιο περιγράφεται η διαδικασία που ακολουθήθηκε ώστε να μοντελοποιηθούν πειραματικά δεδομένα λείανσης μετάλλων με χρήση νευρωνικών δικτύων. Τα πειραματικά δεδομένα ελήφθησαν έτοιμα από την βιβλιογραφία [14], και ακολούθως έγινε απόπειρα μοντελοποίησης τους με χρήση νευρωνικών δικτύων. Δοκιμάστηκαν νευρωνικά δίκτυα πρόσθιας τροφοδότησης καθώς επίσης και νευρωνικά δίκτυα ακτινικής βάσης.

### 4.1 Πειραματικά δεδομένα

Τα πειραματικά δεδομένα που χρησιμοποιήθηκαν, ελήφθησαν έτοιμα από την βιβλιογραφία [14]. Δόθηκαν δεδομένα για τρία διαφορετικά είδη χάλυβα, τα οποία υπέστησαν κατεργασία με έξι διαφορετικούς τύπους τροχών.

Τα πειραματικά δεδομένα που δόθηκαν προς μαθηματική μοντελοποίηση, εμφανίζονται στους πίνακες 5, 6, 7.

Ως δεδομένα εισόδου λαμβάνονται ο τύπος τροχού (τύπος από 1 έως 6), ο τύπος του χάλυβα (1, 2 ή 3) και το βάθος κοπής  $a$  το οποίο λαμβάνει τις τιμές 0.01, 0.02, 0.03, 0.05mm και στα τρία πειράματα.

Το φυσικό μέγεθος που γίνεται απόπειρα να προβλεφθεί για διάφορους τύπους τροχών και για διάφορα βάθη κοπής είναι η τραχύτητα  $R$ .

Προτού γίνει οποιοσδήποτε υπολογισμός, τα δεδομένα είναι απαραίτητο να σταθμιστούν (scaling), ώστε να έχουν την ίδια τάξη μεγέθους. Η στάθμιση των δεδομένων έγινε ώστε ελάχιστη τιμή να είναι η τιμή 0 και μέγιστη η τιμή 1, ακολουθώντας τον παρακάτω τύπο:

$$x_{scaled} = \frac{x_{raw} - x_{min}}{x_{max} - x_{min}}$$

όπου το  $x_{raw}$  είναι το πειραματικό δεδομένο ή η πειραματική μέτρηση όπως ελήφθη από την βιβλιογραφία, το  $x_{scaled}$  είναι η σταθμισμένη του τιμή μεταξύ 0 και 1, και  $x_{min}$ ,  $x_{max}$  οι μέγιστες και ελάχιστες τιμές όπως αυτές εμφανίστηκαν στις πειραματικές μετρήσεις.

Οι πίνακες των δεδομένων αφού πραγματοποιήθηκε η στάθμιση, εμφανίζονται παρακάτω.

<b>Χάλυβας 100Cr6</b>				
A/A Πειράματος	Λειαντικός Τροχός	a (mm)	P1 (W/mm)	R ( $\mu\text{m}$ )
1	1	0.01	93.5	1.68
2		0.02	143.5	1.71
3		0.03	160.0	1.56
4		0.05	278.5	1.43
5	2	0.01	94.0	1.8
6		0.02	164.0	2.07
7		0.03	237.5	1.86
8		0.05	288.0	1.3
9	3	0.01	96.0	1.64
10		0.02	176.0	0.75
11		0.03	208.0	0.75
12		0.05	298.5	0.83
13	4	0.01	99.0	2.07
14		0.02	132.5	1.92
15		0.03	160.0	1.55
16		0.05	256.0	1.69
17	5	0.01	117.5	0.82
18		0.02	170.5	1.03
19		0.03	176.0	1.69
20		0.05	277.5	1.37
21	6	0.01	102.5	1.21
22		0.02	141.0	1.39
23		0.03	202.5	1.3
24		0.05	378.5	2.11

**Πίνακας 5: Πειραματικά δεδομένα για χάλυβα 100Cr6, όπως δόθηκαν από την βιβλιογραφία**

Χάλυβας C45				
A/A Πειράματος	Λειαντικός Τροχός	a (mm)	P1 (W/mm)	R ( $\mu$ m)
25	1	0.01	51.0	1.31
26		0.02	96.5	1.04
27		0.03	109.5	1.54
28		0.05	224.0	1.36
29	2	0.01	48.0	1.8
30		0.02	81.0	1.99
31		0.03	133.5	1.99
32		0.05	194.5	1.9
33	3	0.01	42.5	0.55
34		0.02	90.5	0.96
35		0.03	106.5	0.63
36		0.05	208.0	0.95
37	4	0.01	56.0	1.46
38		0.02	117.5	1.57
39		0.03	140.0	1.78
40		0.05	234.5	1.43
41	5	0.01	45.5	0.92
42		0.02	102.5	0.78
43		0.03	117.5	0.99
44		0.05	216.0	1.08
45	6	0.01	64.0	1.55
46		0.02	103.0	1.53
47		0.03	120.0	1.35
48		0.05	266.5	1.8

**Πίνακας 6: Πειραματικά δεδομένα για χάλυβα C45, όπως προέκυψαν από την βιβλιογραφία**

<b>Χάλυβας X210Cr12</b>				
A/A Πειράματος	Λειαντικός Τροχός	a (mm)	P1 (W/mm)	R ( $\mu\text{m}$ )
49	1	0.01	122.5	0.98
50		0.02	206.0	0.89
51		0.03	340.0	1.32
52		0.05	368.0	1.74
53	2	0.01	141.0	0.67
54		0.02	271.5	0.57
55		0.03	368.0	1.32
56		0.05	426.5	0.94
57	3	0.01	170.5	0.95
58		0.02	330.5	1.31
59		0.03	346.5	1.16
60		0.05	450.5	2.14
61	4	0.01	128.0	0.83
62		0.02	219.5	1.59
63		0.03	352.0	1.67
64		0.05	378.5	1.28
65	5	0.01	138.5	1.4
66		0.02	261.5	1.64
67		0.03	330.5	1.42
68		0.05	421.5	1.25
69	6	0.01	152.5	0.78
70		0.02	241.5	0.89
71		0.03	309.5	0.46
72		0.05	501.5	0.42

**Πίνακας 7: Πειραματικά δεδομένα για χάλυβα X210Cr12, όπως δόθηκαν από την βιβλιογραφία**

Χάλυβας 100Cr6				
	Δεδομένα εισόδου νευρωνικών δικτύων		Δεν χρησιμοποιήθηκαν	Πειραματικά αποτελέσματα
A/A Πειράματος	Λειαντικός Τροχός	a (scaled)	P1 (scaled)	R (scaled)
1	0	0.000	0.000	0.684
2	0	0.250	0.175	0.706
3	0	0.500	0.233	0.596
4	0	1.000	0.649	0.500
5	0.2	0.000	0.002	0.772
6	0.2	0.250	0.247	0.971
7	0.2	0.500	0.505	0.816
8	0.2	1.000	0.682	0.404
9	0.4	0.000	0.009	0.654
10	0.4	0.250	0.289	0.000
11	0.4	0.500	0.402	0.000
12	0.4	1.000	0.719	0.059
13	0.6	0.000	0.019	0.971
14	0.6	0.250	0.137	0.860
15	0.6	0.500	0.233	0.588
16	0.6	1.000	0.570	0.691
17	0.8	0.000	0.084	0.051
18	0.8	0.250	0.270	0.206
19	0.8	0.500	0.289	0.691
20	0.8	1.000	0.646	0.456
21	1	0.000	0.032	0.338
22	1	0.250	0.167	0.471
23	1	0.500	0.382	0.404
24	1	1.000	1.000	1.000

**Πίνακας 8: Σταθμισμένα πειραματικά δεδομένα για χάλυβα 100Cr6**

Χάλυβας C45				
	Δεδομένα εισόδου νευρωνικών δικτύων		Δεν χρησιμοποιήθηκαν	Πειραματικά αποτελέσματα
A/A Πειράματος	Λειαντικός Τροχός	a (scaled)	P1 (scaled)	R (scaled)
25	0	0.000	0.038	0.528
26	0	0.250	0.241	0.340
27	0	0.500	0.299	0.688
28	0	1.000	0.810	0.563
29	0.2	0.000	0.025	0.868
30	0.2	0.250	0.172	1.000
31	0.2	0.500	0.406	1.000
32	0.2	1.000	0.679	0.938
33	0.4	0.000	0.000	0.000
34	0.4	0.250	0.214	0.285
35	0.4	0.500	0.286	0.056
36	0.4	1.000	0.739	0.278
37	0.6	0.000	0.060	0.632
38	0.6	0.250	0.335	0.708
39	0.6	0.500	0.435	0.854
40	0.6	1.000	0.857	0.611
41	0.8	0.000	0.013	0.257
42	0.8	0.250	0.268	0.160
43	0.8	0.500	0.335	0.306
44	0.8	1.000	0.775	0.368
45	1	0.000	0.096	0.694
46	1	0.250	0.270	0.681
47	1	0.500	0.346	0.556
48	1	1.000	1.000	0.868

**Πίνακας 9:** Σταθμισμένα πειραματικά δεδομένα για χάλυβα C45

Χάλυβας X210Cr12				
	Δεδομένα εισόδου νευρωνικών δικτύων		Δεν χρησιμοποιήθηκαν	Πειραματικά αποτελέσματα
A/A Πειράματος	Λειαντικός Τροχός	a (scaled)	P1 (scaled)	R (scaled)
49	0	0.000	0.000	0.326
50	0	0.250	0.220	0.273
51	0	0.500	0.574	0.523
52	0	1.000	0.648	0.767
53	0.2	0.000	0.049	0.145
54	0.2	0.250	0.393	0.087
55	0.2	0.500	0.648	0.523
56	0.2	1.000	0.802	0.302
57	0.4	0.000	0.127	0.308
58	0.4	0.250	0.549	0.517
59	0.4	0.500	0.591	0.430
60	0.4	1.000	0.865	1.000
61	0.6	0.000	0.015	0.238
62	0.6	0.250	0.256	0.680
63	0.6	0.500	0.606	0.727
64	0.6	1.000	0.675	0.500
65	0.8	0.000	0.042	0.570
66	0.8	0.250	0.367	0.709
67	0.8	0.500	0.549	0.581
68	0.8	1.000	0.789	0.483
69	1	0.000	0.079	0.209
70	1	0.250	0.314	0.273
71	1	0.500	0.493	0.023
72	1	1.000	1.000	0.000

**Πίνακας 10: Σταθμισμένα πειραματικά δεδομένα για χάλυβα X210Cr12**

Ο τύπος του λειαντικού τροχού και το βάθος κοπής, αφού σταθμίστηκαν, χρησιμοποιήθηκαν ως δεδομένα εισόδου των νευρωνικών δικτύων, ενώ τα πειραματικά αποτελέσματα τραχύτητας  $R$ , ήταν τα δεδομένα με τα οποία έγινε η εκπαίδευση και ο έλεγχος των νευρωνικών δικτύων που κατασκευάστηκαν.

## 4.2 Κατασκευή νευρωνικών δικτύων πρόσθιας τροφοδότησης

Κατά την απόπειρα μαθηματικής μοντελοποίησης των πειραματικών δεδομένων δοκιμάστηκε ένας μεγάλος αριθμός διαφορετικών νευρωνικών δικτύων πρόσθιας τροφοδότησης με μία και δύο κρυμμένες στοιβάδες νευρώνων.

Τα νευρωνικά δίκτυα μίας κρυμμένης στοιβάδας νευρώνων είχαν από 4 έως 10 νευρώνες στην κρυμμένη τους στοιβάδα.

Τα νευρωνικά δίκτυα με δύο κρυμμένες στοιβάδες νευρώνων επίσης είχαν από 4 έως 10 νευρώνες σε κάθε στοιβάδα. Δοκιμάστηκαν όλοι οι πιθανοί συνδυασμοί αριθμού νευρώνων μεταξύ των στοιβάδων, οι οποίοι είχαν διαφορά έως και 2 νευρώνων μεταξύ των στοιβάδων. Ήτοι, π.χ. αν η πρώτη στοιβάδα νευρώνων είχε 7 νευρώνες, έγιναν δοκιμές με την δεύτερη στοιβάδα να έχει 5, 6, 7, 8, 9 νευρώνες, κ.ο.κ.

Για κάθε νευρωνικό δίκτυο έγιναν τρεις δοκιμές με διαφορετική αρχικοποίηση των βαρών και βρέθηκε το βέλτιστο δίκτυο που προσέγγιζε τα δεδομένα με το μικρότερο δυνατό σφάλμα και το μέγιστο δυνατό συντελεστή συσχέτισης R.

Για το βέλτιστο δίκτυο, δοκιμάστηκαν επίσης και εναλλακτικοί αλγόριθμοι εκπαίδευσης.

Όλες οι δοκιμές έγιναν με χρήση του MATLAB, χρησιμοποιώντας τις έτοιμες συναρτήσεις και τα εργαλεία νευρωνικών δικτύων. Οι εντολές που χρησιμοποιήθηκαν είναι οι εξής:

- Για δημιουργία νευρωνικού δικτύου μίας κρυμμένης στοιβάδας με π.χ. 5 νευρώνες:  
`net = feedforwardnet(5,'trainlm')`
- Για δημιουργία νευρωνικού δικτύου δύο κρυμμένων στοιβάδων με π.χ. 7 νευρώνες η πρώτη και 5 νευρώνες η δεύτερη:  
`net = feedforwardnet(7, 5,'trainlm')`

Ο όρος "trainlm" στην παραπάνω εντολή επιλέγει τον αλγόριθμο εκπαίδευσης Levenberg-Marquardt ο οποίος ήταν ο κύριος αλγόριθμος που χρησιμοποιήθηκε.



Εκτός από τον αλγόριθμο Levenberg-Marquardt, δοκιμάστηκαν και οι παρακάτω αλγόριθμοι, όπως υπάρχουν στο περιβάλλον του MATLAB:

- "Conjugate Gradient backpropagation, with Powell-Beale restarts", χρησιμοποιώντας την εντολή "traincgb"
- "Conjugate Gradient backpropagation, with Fletcher-Reeves updates", χρησιμοποιώντας την εντολή "traincgf"
- "Conjugate Gradient backpropagation, with Polac-Ribiere updates", χρησιμοποιώντας την εντολή "traincgp"

Και οι τέσσερις αλγόριθμοι που δοκιμάστηκαν είναι παραλλαγές αλγορίθμων επιτηρούμενης εκπαίδευσης. Βασίζονται στην μεταβολή των βαρών προς την κατεύθυνση της μερικής παραγώγου του σφάλματος, όπως εξηγήθηκε στο κεφάλαιο 2.

Όπως αναλύθηκε και στο κεφάλαιο 2, η διαδικασία εκπαίδευσης νευρωνικών δικτύων πρόσθιας τροφοδότησης, περιλαμβάνει προσδιορισμό των βαρών  $w_{ij}$  των νευρωνικών δικτύων σε τρία στάδια.

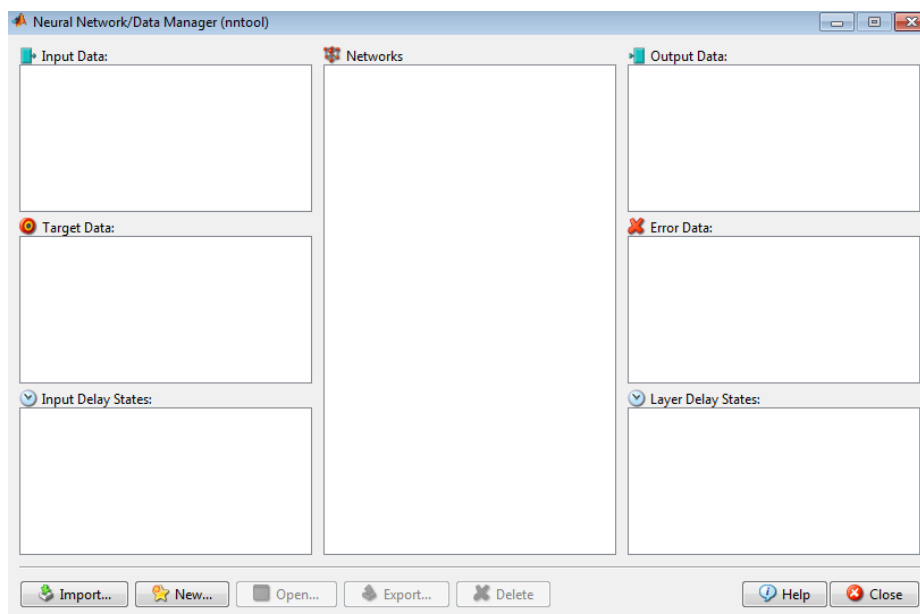
- Ορισμός βαρών με τυχαίο τρόπο
- Υπολογισμός σφάλματος
- Μεταβολή του κάθε βάρους προς την αντίθετη κατεύθυνση της μερικής παραγώγου του σφάλματος ως προς το βάρος αυτό

Το MATLAB διαχωρίζει τα δεδομένα σε τρεις ομάδες:

- Train: Χρησιμοποιούνται για την εκπαίδευση των νευρωνικών δικτύων
- Validation: Χρησιμοποιούνται για την αξιολόγηση του νευρωνικού δικτύου με ένα ξεχωριστό δείγμα δεδομένων που δεν χρησιμοποιήθηκε κατά την εκπαίδευση.
- Test: Χρησιμοποιούνται για την τελική αξιολόγηση του νευρωνικού δικτύου.

Τα προγράμματα που κατασκευάστηκαν στο MATLAB, για να γίνει η μελέτη, περιελάμβαναν τις παραπάνω εντολές και συνέβαλλαν στην αυτοματοποίηση των δοκιμών που πραγματοποιήθηκαν.

Εκτός από τα scripts, χρησιμοποιήθηκε και το nntool (neural network tool) του MATLAB (βλ. σχήμα 33).



Σχήμα 33: Εργαλειοθήκη νευρωνικών δικτύων MATLAB

### 4.3 Κατασκευή νευρωνικών δικτύων ακτινικής βάσης

Εκτός από τα νευρωνικά δίκτυα πρόσθιας τροφοδότησης, δοκιμάστηκαν επίσης και νευρωνικά δίκτυα ακτινικής βάσης. Όπως συζητήθηκε στο κεφάλαιο 3, τα νευρωνικά δίκτυα ακτινικής βάσης έχουν μόνον μία κρυμμένη στοιβάδα νευρώνων. Δοκιμάστηκαν νευρωνικά δίκτυα ακτινικής βάσης με 5, 10, 15, ..., 60 νευρώνες στην κρυφή τους στοιβάδα.

Η συνάρτηση ενεργοποίησης των νευρωνικών δικτύων ακτινικής βάσης που χρησιμοποιήθηκε ήταν η συνάρτηση Gauss. Για κάθε νευρωνικό δίκτυο ακτινικής βάσης, δοκιμάστηκαν διαφορετικές τιμές spread της συνάρτησης Gauss. Συγκεκριμένα, δοκιμάστηκαν οι τιμές 0.25, 0.5, 1, 2.

Τα νευρωνικά δίκτυα ακτινικής βάσης κατασκευάστηκαν στο MATLAB με χρήση των έτοιμων βιβλιοθηκών του. Η εντολή για να δημιουργηθεί ένα νευρωνικό δίκτυο ακτινικής βάσης π.χ. με spread 0.5 και αριθμό νευρώνων κρυμμένης στοιβάδας ίσο με 10, στο MATLAB, είναι η εξής:

- newrb(I, O, 0.5, 10)

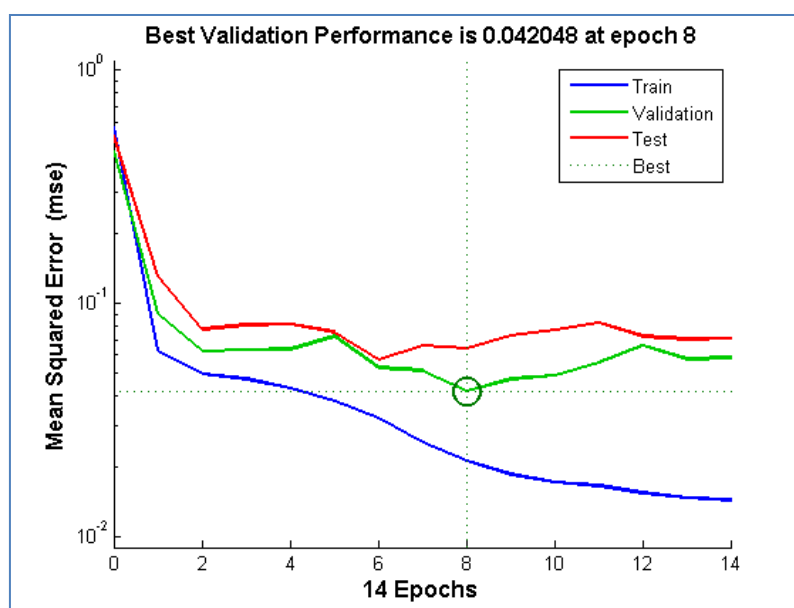
Όπως και με τα νευρωνικά δίκτυα πρόσθιας τροφοδότησης, έτσι και στα νευρωνικά δίκτυα ακτινικής βάσης, κατασκευάστηκαν scripts, τα οποία αυτοματοποίησαν την διαδικασία που περιγράφηκε παραπάνω.

## 5. Αποτελέσματα

Στο παρόν κεφάλαιο παρουσιάζονται τα αποτελέσματα που προέκυψαν ακολουθώντας τις διαδικασίες υπολογισμών που αναλύθηκαν στο κεφάλαιο 4.

### 5.1 Αποτελέσματα νευρωνικών δικτύων πρόσθιας τροφοδότησης

Όπως προαναφέρθηκε, δοκιμάστηκε ένας μεγάλος αριθμός από αρχιτεκτονικές νευρωνικών δικτύων. Κάθε αρχιτεκτονική εκπαιδεύτηκε και δοκιμάστηκε τρεις φορές στο MATLAB για να υπάρξει διαφορετική αρχικοποίηση. Το βέλτιστο νευρωνικό δίκτυο, με βάση το ελάχιστο σφάλμα ( $MSE = \text{Mean Squared Error}$ ) είχε μία κρυμμένη στοιβάδα με έξι νευρώνες. Το διάγραμμα εκπαίδευσης του παρουσιάζεται στο σχήμα 34.

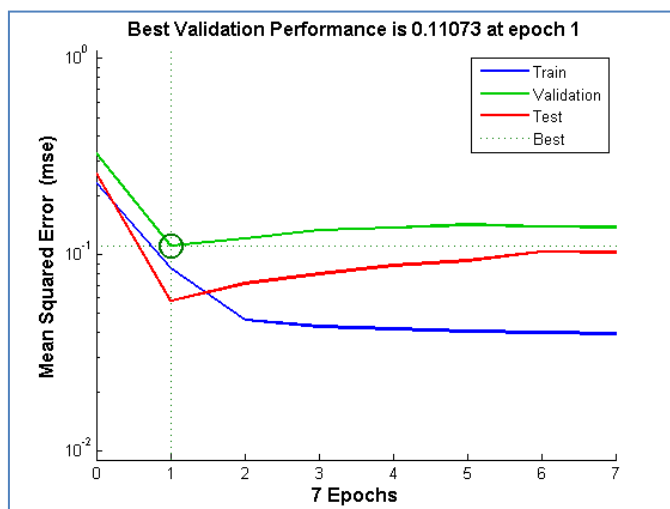


Σχήμα 34: Διάγραμμα εκπαίδευσης βέλτιστου δικτύου με μία κρυμμένη στοιβάδα 6 νευρώνων

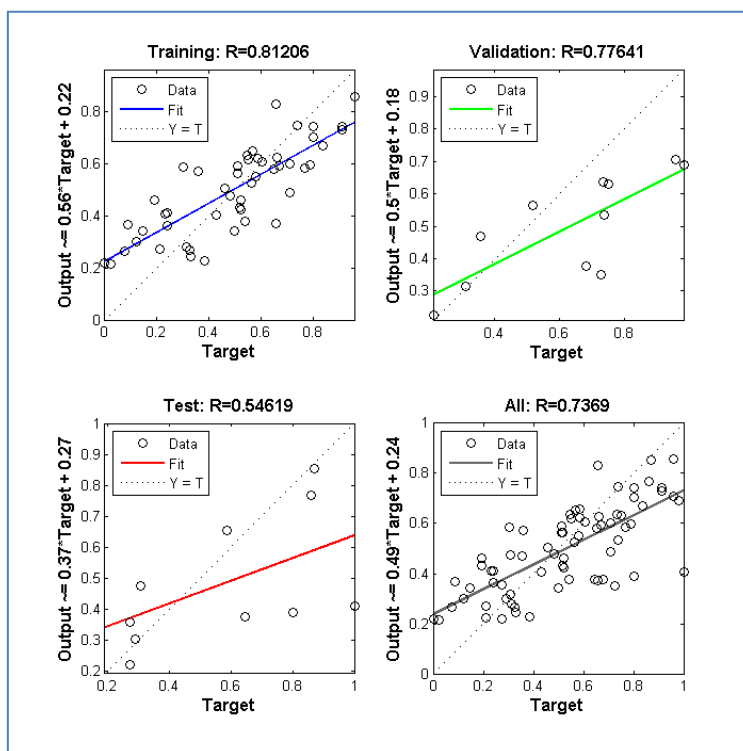
Αντίστοιχα, το διάγραμμα εκπαίδευσης ενός νευρωνικού δικτύου με κατώτερα αποτελέσματα (2 κρυμμένες στοιβάδες, 4 νευρώνες η κάθε στοιβάδα), παρουσιάζεται στο σχήμα 35.

Από τα διαγράμματα 34, 35 είναι δυνατό να φανεί ότι το νευρωνικό δίκτυο με μία κρυμμένη στοιβάδα και 6 νευρώνες σε αυτήν, ήταν σε θέση να μειώνει συνεχώς το σφάλμα εκπαίδευσης, καθ' όλη τη διάρκειά της.

Στο σχήμα 36, παρουσιάζεται η διασπορά των προβλέψεων του βέλτιστου νευρωνικού δικτύου, σε σχέση με τα πειραματικά δεδομένα.

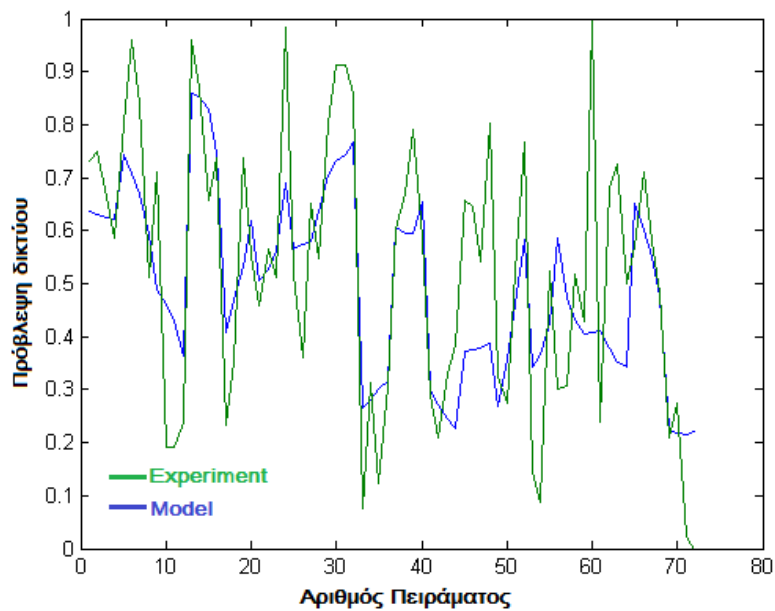


Σχήμα 35: Διάγραμμα εκπαίδευσης νευρωνικού δικτύου με δύο κρυμμένες στοιβάδες με 4 νευρώνες έκαστη.



Σχήμα 36: Διασπορά των πειραματικών δεδομένων σε σχέση με τις προβλέψεις του βέλτιστου νευρωνικού δικτύου

Η ικανότητα πρόβλεψης του βέλτιστου νευρωνικού δικτύου είναι φανερή στο σχήμα 37, όπου απεικονίζονται οι πειραματικές τιμές και προβλέψεις δικτύου (άξονας y) σε σχέση με τον αριθμό πειράματος (άξονας x).



**Σχήμα 37: Προβλέψεις πειραματικών δεδομένων βέλτιστου νευρωνικού δικτύου.**

Τα αποτελέσματα των νευρωνικών δικτύων πρόσθιας τροφοδότησης συνοψίζονται στους πίνακες 11, 12, 13 για τις δοκιμές που έγιναν σε νευρωνικά δίκτυα μίας κρυμμένης στοιβάδας, και στους πίνακες 14, 15, 16 για τις δοκιμές που έγιναν σε νευρωνικά δίκτυα δύο κρυμμένων στοιβάδων.

Αριθμός Νευρώνων κρυμμένης στοιβάδας	MSE (training)	MSE (validation)	MSE (testing)	MSE (ολικό)	R (training)	R (validation)	R (test)	R (ολικό)
4	4.71E-02	4.87E-02	7.64E-02	1.72E-01	2.98E-01	5.30E-01	1.90E-01	3.03E-01
5	2.51E-02	7.12E-02	1.16E-01	2.12E-01	6.20E-01	-2.19E-01	-3.79E-01	3.63E-01
6	1.43E-02	4.20E-02	5.71E-02	1.13E-01	8.12E-01	7.76E-01	5.46E-01	7.37E-01
7	3.14E-02	6.73E-02	5.96E-02	1.58E-01	5.28E-01	9.56E-02	2.38E-01	4.33E-01
8	2.73E-02	5.73E-02	1.02E-01	1.86E-01	5.21E-01	2.95E-01	4.05E-01	4.43E-01
9	1.85E-02	2.69E-02	1.24E-01	1.70E-01	7.80E-01	5.23E-01	-1.99E-01	4.78E-01
10	2.22E-02	1.02E-01	9.15E-02	2.15E-01	6.41E-01	2.12E-01	1.95E-01	4.47E-01

**Πίνακας 11: Πίνακας αποτελεσμάτων για νευρωνικά δίκτυα πρόσθιας τροφοδότησης, μίας κρυμμένης στοιβάδας (Δοκιμή 1)**

Αριθμός Νευρώνων κρυμμένης στοιβάδας	MSE (training)	MSE (validation)	MSE (testing)	MSE (ολικό)	R (training)	R (validation)	R (test)	R (ολικό)
4	4.28E-02	3.60E-02	6.04E-02	1.39E-01	3.95E-01	6.21E-01	5.59E-02	3.27E-01
5	4.75E-02	3.79E-02	3.99E-02	1.25E-01	3.02E-01	7.39E-01	4.13E-01	3.79E-01
6	2.86E-02	5.16E-02	1.09E-01	1.89E-01	6.66E-01	3.74E-01	-4.71E-01	4.52E-01
7	2.46E-02	4.13E-02	1.16E-01	1.82E-01	7.32E-01	5.90E-01	1.80E-01	5.83E-01
8	2.22E-02	1.91E-02	5.70E-02	9.83E-02	8.16E-01	8.62E-01	-8.67E-02	6.84E-01
9	2.19E-02	7.06E-02	5.08E-02	1.43E-01	3.48E-01	8.26E-01	4.08E-01	4.12E-01
10	2.51E-02	8.21E-02	7.88E-02	1.86E-01	1.75E-01	-1.35E-01	4.38E-02	1.23E-01

**Πίνακας 12: Πίνακας αποτελεσμάτων για νευρωνικά δίκτυα πρόσθιας τροφοδότησης, μίας κρυμμένης στοιβάδας (Δοκιμή 2)**

Αριθμός Νευρώνων κρυμμένης στοιβάδας	MSE (training)	MSE (validation)	MSE (testing)	MSE (ολικό)	R (training)	R (validation)	R (test)	R (ολικό)
4	4.05E-02	4.94E-02	4.96E-02	1.39E-01	-2.60E-01	6.73E-01	-4.71E-01	-2.37E-01
5	2.46E-02	7.14E-02	6.78E-02	1.64E-01	4.04E-01	-1.00E-01	-3.62E-01	2.37E-01
6	3.09E-02	7.21E-02	5.89E-02	1.62E-01	4.73E-01	4.06E-01	2.92E-01	4.15E-01
7	3.56E-02	8.25E-02	5.37E-02	1.72E-01	2.75E-01	4.72E-01	-1.13E-01	2.91E-01
8	2.30E-02	5.53E-02	1.17E-01	1.95E-01	5.00E-01	1.41E-01	-5.21E-01	2.91E-01
9	3.09E-02	5.65E-02	6.08E-02	1.48E-01	4.60E-01	2.61E-01	1.25E-01	3.84E-01
10	2.55E-02	9.98E-02	4.71E-02	1.72E-01	6.53E-01	4.41E-01	2.07E-01	4.94E-01

**Πίνακας 13: Πίνακας αποτελεσμάτων για νευρωνικά δίκτυα πρόσθιας τροφοδότησης, μίας κρυμμένης στοιβάδας (Δοκιμή 3)**

Καθώς σε κάθε δοκιμή γίνεται διαφορετική αρχικοποίηση των βαρών, παράγονται διαφορετικά αποτελέσματα, ακόμα και αν η αρχιτεκτονική των

νευρωνικών δικτύων είναι ίδια. Το ίδιο παρατηρείται και στους υπολογισμούς με δύο κρυμμένες στοιβάδες (πίνακες 14, 15 και 16).

Αριθμός Νευρώνων κρυμμένης στοιβάδας 1	Αριθμός Νευρώνων κρυμμένης στοιβάδας 2	MSE (training)	MSE (validation)	MSE (testing)	MSE (ολικό)	R (training)	R (validation)	R (test)	R (ολικό)
4	4	3.94E-02	1.11E-01	5.82E-02	2.08E-01	2.30E-01	4.19E-01	4.16E-01	2.86E-01
4	5	4.05E-02	7.85E-02	5.70E-02	1.76E-01	3.69E-01	3.13E-01	2.01E-01	2.22E-01
5	4	3.73E-02	4.87E-02	3.89E-02	1.25E-01	4.11E-01	2.71E-01	4.52E-01	3.98E-01
4	6	3.76E-02	4.65E-02	6.69E-02	1.51E-01	4.64E-01	5.64E-02	2.82E-01	3.89E-01
5	5	2.20E-02	9.08E-02	9.83E-02	2.11E-01	6.44E-01	-3.37E-01	1.96E-02	3.87E-01
6	4	2.91E-02	3.66E-02	6.70E-02	1.33E-01	4.81E-01	4.81E-01	4.37E-01	4.72E-01
5	6	3.89E-02	5.32E-02	5.27E-02	1.45E-01	4.67E-01	1.11E-01	1.52E-01	3.89E-01
6	5	3.03E-02	4.89E-02	3.77E-02	1.17E-01	6.49E-01	6.88E-01	-5.52E-01	5.05E-01
5	7	2.24E-02	1.34E-01	1.23E-01	2.79E-01	4.91E-01	2.41E-01	-6.05E-02	2.82E-01
6	6	2.88E-02	8.55E-02	6.84E-02	1.83E-01	8.58E-02	6.40E-01	-9.05E-02	1.64E-01
7	5	3.32E-02	9.79E-02	6.19E-02	1.93E-01	2.30E-01	-1.14E-01	-1.28E-01	1.08E-01
6	7	3.60E-02	5.48E-02	6.79E-02	1.59E-01	1.78E-01	4.56E-01	1.18E-01	2.11E-01
7	6	1.92E-02	8.61E-02	7.39E-02	1.79E-01	4.88E-01	-1.09E-01	2.62E-01	3.67E-01
6	8	1.50E-02	8.36E-02	4.82E-02	1.47E-01	4.03E-01	2.28E-01	2.87E-01	3.41E-01
7	7	1.42E-02	7.41E-02	7.91E-02	1.67E-01	5.83E-01	4.34E-01	8.90E-02	4.48E-01
8	6	3.54E-02	4.64E-02	5.30E-02	1.35E-01	1.26E-01	7.57E-02	-6.43E-01	4.90E-02
7	8	3.83E-03	4.22E-02	1.17E-01	1.63E-01	7.03E-01	3.23E-01	5.35E-02	5.16E-01
8	7	3.97E-03	7.76E-02	7.05E-02	1.52E-01	7.08E-01	2.24E-01	-1.38E-01	5.36E-01
7	9	2.33E-02	8.07E-02	2.92E-02	1.33E-01	6.02E-01	6.52E-02	-3.86E-01	4.85E-01
8	8	2.35E-02	7.73E-02	2.43E-02	1.25E-01	3.09E-01	1.61E-01	3.79E-01	2.88E-01
9	7	1.20E-02	5.35E-02	5.61E-02	1.22E-01	2.44E-01	-2.66E-02	3.58E-01	2.08E-01
8	9	8.38E-03	1.35E-01	8.57E-02	2.29E-01	2.25E-01	-5.55E-01	3.59E-01	1.06E-01
9	8	2.15E-02	3.82E-02	5.20E-02	1.12E-01	8.29E-02	4.33E-01	-6.57E-02	1.14E-01
8	10	1.21E-02	7.53E-02	1.17E-01	2.05E-01	6.37E-01	2.10E-01	-5.49E-01	3.93E-01
9	9	2.35E-02	9.63E-02	4.14E-02	1.61E-01	5.47E-01	-6.63E-02	3.96E-01	4.41E-01
10	8	6.77E-03	6.83E-02	6.10E-02	1.36E-01	6.41E-01	-1.08E-01	3.16E-03	4.26E-01
9	10	2.01E-02	7.97E-02	4.22E-02	1.42E-01	5.80E-01	3.37E-01	6.84E-01	5.29E-01
10	9	5.36E-03	1.00E-01	5.07E-02	1.57E-01	6.97E-01	1.37E-01	4.41E-03	4.87E-01
10	10	1.02E-02	1.18E-01	5.58E-02	1.84E-01	8.44E-02	3.83E-01	4.13E-01	1.50E-01

**Πίνακας 14: Πίνακας αποτελεσμάτων για νευρωνικά δίκτυα δύο κρυμμένων στοιβάδων (Δοκιμή 1)**



Αριθμός Νευρώνων κρυμμένης στοιβάδας 1	Αριθμός Νευρώνων κρυμμένης στοιβάδας 2	MSE (training)	MSE (validation)	MSE (testing)	MSE (ολικό)	R (training)	R (validation)	R (test)	R (ολικό)
4	4	2.48E-02	1.01E-01	6.01E-02	1.86E-01	5.35E-01	2.93E-01	1.58E-01	3.87E-01
4	5	2.91E-02	6.96E-02	5.16E-02	1.50E-01	3.50E-01	5.10E-01	3.15E-02	3.29E-01
5	4	2.31E-02	5.39E-02	6.50E-02	1.42E-01	4.72E-01	3.41E-01	-4.46E-01	3.31E-01
4	6	3.81E-02	8.91E-02	6.76E-02	1.95E-01	3.74E-01	4.37E-01	-7.25E-03	3.35E-01
5	5	1.60E-02	2.15E-02	6.36E-02	1.01E-01	6.47E-01	3.39E-01	4.02E-01	5.99E-01
6	4	3.51E-02	7.62E-02	7.20E-02	1.83E-01	-1.00E-01	5.21E-01	-9.22E-02	-1.60E-02
5	6	1.04E-02	8.23E-02	1.07E-01	1.99E-01	6.09E-01	1.25E-01	-2.00E-01	3.75E-01
6	5	2.10E-02	7.08E-02	4.71E-02	1.39E-01	4.89E-01	-1.66E-02	6.12E-01	4.53E-01
5	7	3.11E-02	2.64E-02	6.82E-02	1.26E-01	5.68E-01	6.78E-01	8.86E-02	5.06E-01
6	6	2.54E-02	9.96E-02	8.81E-02	2.13E-01	5.54E-01	-3.95E-02	2.08E-01	4.05E-01
7	5	2.32E-02	6.49E-02	7.11E-02	1.59E-01	3.54E-01	4.29E-01	-7.25E-03	3.11E-01
6	7	1.69E-02	7.01E-02	2.88E-02	1.16E-01	1.43E-01	-1.93E-01	3.67E-01	1.09E-01
7	6	2.08E-02	9.82E-02	6.76E-02	1.87E-01	6.53E-01	3.44E-01	-2.13E-01	4.26E-01
6	8	1.07E-02	1.23E-01	7.12E-02	2.05E-01	6.88E-01	-3.18E-01	1.73E-01	4.82E-01
7	7	3.00E-02	5.40E-02	3.78E-02	1.22E-01	4.83E-01	2.89E-01	5.65E-01	4.63E-01
8	6	3.24E-02	5.46E-02	1.13E-01	2.00E-01	1.74E-01	3.34E-01	-2.79E-01	1.19E-01
7	8	4.05E-02	7.88E-02	5.63E-02	1.76E-01	1.50E-01	7.97E-02	4.47E-01	1.16E-01
8	7	7.15E-04	1.18E-01	9.94E-02	2.18E-01	9.63E-01	2.96E-01	3.53E-01	6.66E-01
7	9	1.28E-02	8.28E-02	5.52E-02	1.51E-01	7.34E-01	2.84E-01	-2.26E-01	5.47E-01
8	8	2.76E-02	3.29E-02	7.52E-02	1.36E-01	5.30E-01	5.50E-01	3.85E-01	4.73E-01
9	7	6.25E-03	8.93E-02	2.44E-02	1.20E-01	7.54E-01	3.65E-01	7.64E-01	6.67E-01
8	9	2.74E-02	7.98E-02	1.18E-01	2.25E-01	6.10E-01	3.99E-01	-5.82E-01	3.86E-01
9	8	2.80E-03	1.05E-01	5.84E-02	1.66E-01	6.23E-02	3.18E-02	3.40E-01	4.95E-01
8	10	2.87E-02	4.49E-02	1.11E-01	1.85E-01	-1.14E-01	4.22E-01	1.86E-01	-4.09E-02
9	9	8.31E-03	6.90E-02	9.82E-02	1.76E-01	2.84E-01	1.20E-02	1.05E-01	2.16E-01
10	8	2.13E-02	7.07E-02	6.19E-02	1.54E-01	3.12E-01	6.22E-01	-1.05E-01	3.03E-01
9	10	4.76E-03	7.38E-02	7.91E-02	1.58E-01	4.27E-01	3.65E-01	-2.07E-01	3.39E-01
10	9	3.79E-04	2.16E-01	1.17E-01	3.33E-01	8.56E-01	1.78E-01	2.11E-01	5.09E-01
10	10	1.41E-02	9.06E-02	6.37E-02	1.68E-01	5.51E-01	2.12E-02	1.32E-01	4.23E-01

**Πίνακας 15: Πίνακας αποτελεσμάτων για νευρωνικά δίκτυα δύο κρυμμένων στοιβάδων (Δοκιμή 2)**

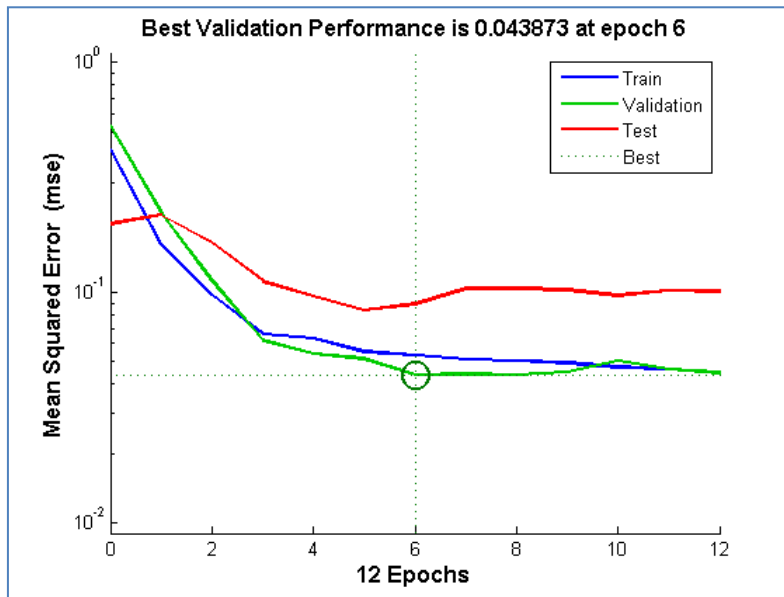
Λόγω του γεγονότος ότι για κάθε αρχικοποίηση λαμβάνονται διαφορετικά αποτελέσματα, το βέλτιστο νευρωνικό δίκτυο επελέγη με βάση α) την σύγκριση σφάλματος και συντελεστή συσχέτισης μεταξύ διαφορετικών αρχιτεκτονικών και β) την σύγκριση μεταξύ των διαφορετικών δοκιμών που έγιναν.

Αριθμός Νευρώνων κρυμμένης στοιβάδας 1	Αριθμός Νευρώνων κρυμμένης στοιβάδας 2	MSE (training)	MSE (validation)	MSE (testing)	MSE (ολικό)	R (training)	R (validation)	R (test)	R (ολικό)
4	4	5.28E-02	7.59E-02	1.52E-02	1.44E-01	5.61E-02	1.77E-01	1.50E-01	1.26E-01
4	5	3.24E-02	9.21E-02	3.41E-02	1.59E-01	4.15E-01	2.16E-01	1.14E-01	3.33E-01
5	4	3.60E-02	4.43E-02	7.50E-02	1.55E-01	2.99E-01	3.19E-01	-1.20E-01	2.31E-01
4	6	4.59E-02	1.00E-01	2.51E-02	1.71E-01	3.92E-01	1.43E-01	3.06E-01	3.41E-01
5	5	1.56E-02	4.54E-02	6.78E-02	1.29E-01	7.13E-01	3.43E-01	6.27E-01	6.38E-01
6	4	2.67E-02	9.20E-02	9.13E-02	2.10E-01	3.76E-01	-4.82E-01	-9.67E-02	1.50E-01
5	6	3.29E-02	2.85E-02	7.61E-02	1.38E-01	2.11E-01	6.99E-01	4.75E-01	2.61E-01
6	5	2.80E-02	5.52E-02	5.60E-02	1.39E-01	6.07E-01	2.90E-01	1.24E-01	4.79E-01
5	7	2.70E-02	7.89E-02	3.78E-02	1.44E-01	5.43E-01	4.18E-01	1.62E-01	4.75E-01
6	6	1.19E-02	8.20E-02	1.34E-01	2.28E-01	7.08E-01	2.68E-01	-2.51E-01	4.72E-01
7	5	3.80E-02	3.03E-02	1.10E-01	1.79E-01	-2.96E-01	5.45E-01	-3.79E-01	-2.01E-01
6	7	3.07E-03	2.76E-02	1.50E-01	1.81E-01	6.64E-01	7.06E-01	1.07E-01	5.16E-01
7	6	2.25E-02	7.99E-02	5.61E-02	1.58E-01	2.24E-01	2.67E-01	2.29E-01	2.37E-01
6	8	3.74E-03	6.90E-02	5.96E-02	1.32E-01	7.72E-01	3.03E-01	1.10E-01	5.75E-01
7	7	2.26E-02	4.98E-02	1.10E-01	1.82E-01	2.56E-01	-1.16E-01	-2.62E-01	1.29E-01
8	6	1.80E-02	4.85E-02	5.81E-02	1.25E-01	-1.49E-01	2.85E-01	-2.96E-01	-9.75E-02
7	8	1.07E-02	6.14E-02	8.30E-02	1.55E-01	7.47E-01	-3.63E-01	-5.35E-01	4.74E-01
8	7	2.05E-02	7.57E-02	1.25E-01	2.21E-01	3.25E-01	2.15E-01	-9.58E-02	2.54E-01
7	9	2.23E-02	4.94E-02	3.54E-02	1.07E-01	6.29E-01	7.08E-01	5.37E-01	6.02E-01
8	8	1.41E-02	7.80E-02	1.21E-01	2.13E-01	3.95E-01	1.82E-01	-4.82E-01	2.43E-01
9	7	1.73E-02	7.46E-02	3.82E-02	1.30E-01	9.36E-02	3.49E-01	3.15E-01	1.86E-01
8	9	1.33E-02	4.93E-02	5.41E-02	1.17E-01	5.36E-01	4.86E-01	2.60E-01	4.76E-01
9	8	5.12E-03	1.11E-01	5.27E-02	1.69E-01	2.95E-01	-2.88E-01	-1.35E-01	1.40E-01
8	10	1.75E-02	4.95E-02	6.84E-02	1.35E-01	5.97E-01	4.92E-01	6.27E-01	5.86E-01
9	9	2.49E-02	8.15E-02	1.21E-01	2.28E-01	3.73E-01	-2.78E-01	-2.93E+04	2.07E-01
10	8	2.56E-02	8.24E-02	8.91E-02	1.97E-01	4.48E-01	-2.64E-02	-5.63E-01	1.81E-01
9	10	2.36E-02	6.20E-02	7.90E-02	1.65E-01	4.45E-01	3.02E-01	2.45E-01	3.77E-01
10	9	1.54E-02	4.76E-02	6.94E-02	1.32E-01	3.74E-01	3.13E-01	-4.20E-02	2.76E-01
10	10	8.80E-03	1.05E-01	1.02E-01	2.15E-01	5.08E-01	1.87E-01	-3.66E-01	3.08E-01

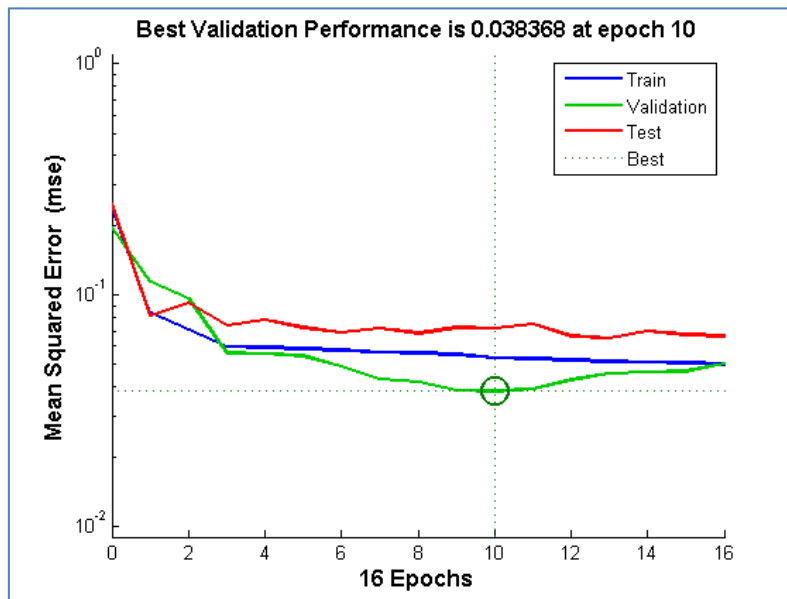
Πίνακας 16: Πίνακας αποτελεσμάτων για νευρωνικά δίκτυα δύο κρυμμένων στοιβάδων (Δοκιμή 3)

### 5.1.1 Επίδραση Αλγορίθμου εκπαίδευσης

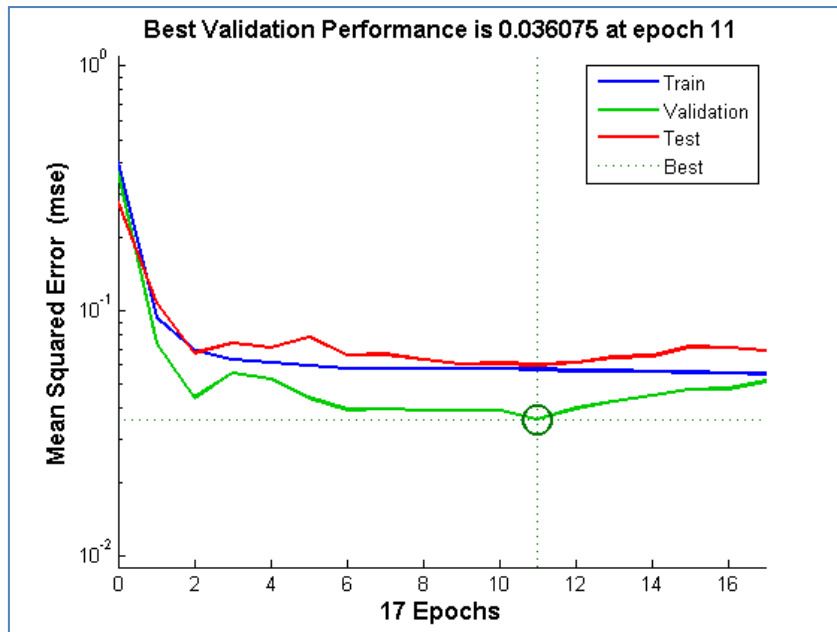
Εκτός από τον αλγόριθμο Levenberg Marquardt, δοκιμάστηκαν οι εναλλακτικοί αλγόριθμοι που περιγράφηκαν στο κεφάλαιο 4. Τα αποτελέσματα εκπαίδευσης που προέκυψαν παρατίθενται στα σχήματα 38 έως 40.



Σχήμα 38: Διάγραμμα σφάλματος κατά την εκπαίδευση νευρωνικού δικτύου μίας κρυμμένης στοιβάδας και 6 νευρώνων με τον αλγόριθμο CGB (βλ. κεφάλαιο 4)

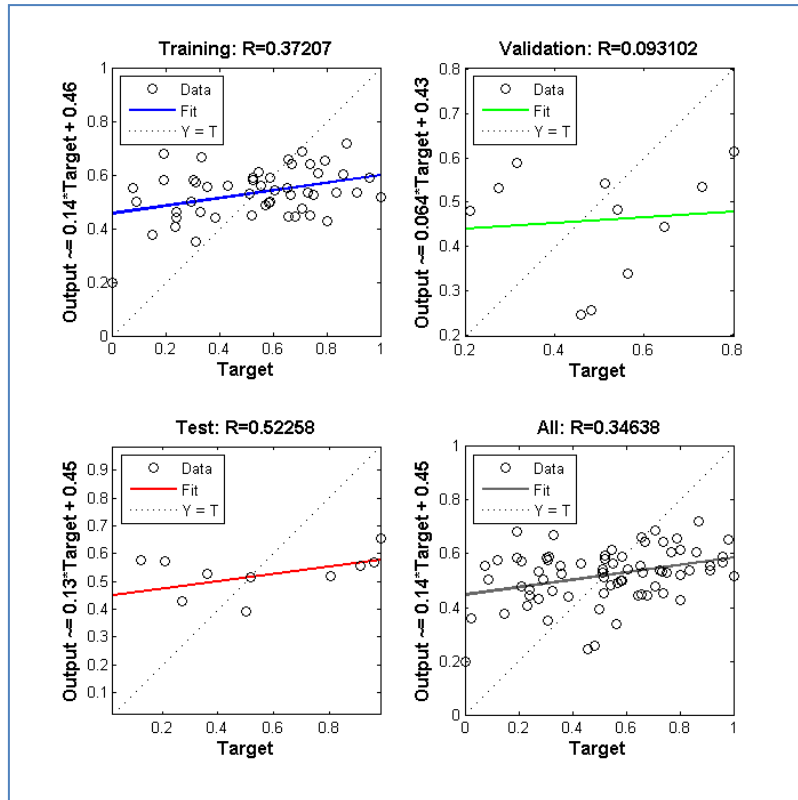


Σχήμα 39: Διάγραμμα σφάλματος κατά την εκπαίδευση νευρωνικού δικτύου μίας κρυμμένης στοιβάδας και 6 νευρώνων με τον αλγόριθμο CGF (βλ. κεφάλαιο 4)

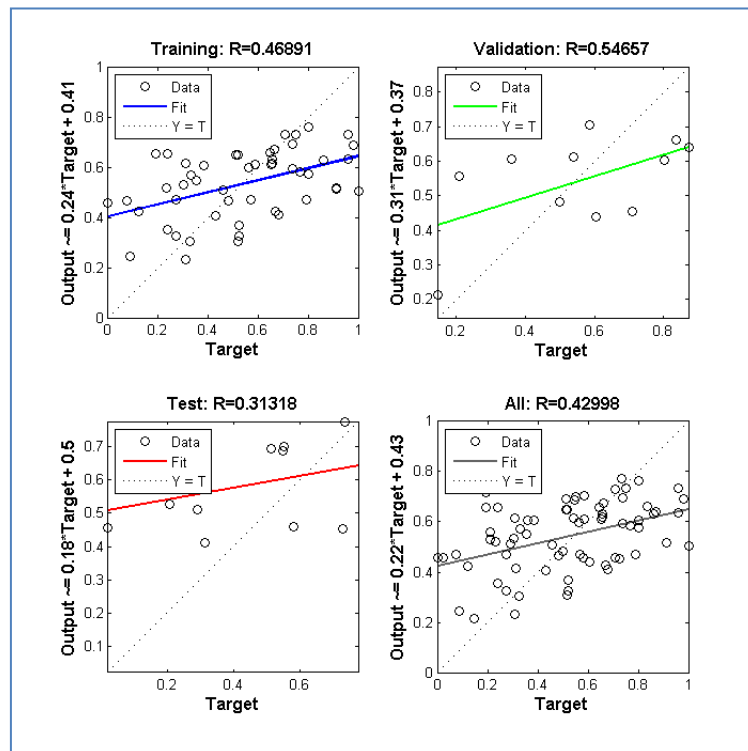


Σχήμα 40: Διάγραμμα σφάλματος κατά την εκπαίδευση νευρωνικού δικτύου μίας κρυμμένης στοιβάδας και 6 νευρώνων με τον αλγόριθμο CGP (βλ. κεφάλαιο 4)

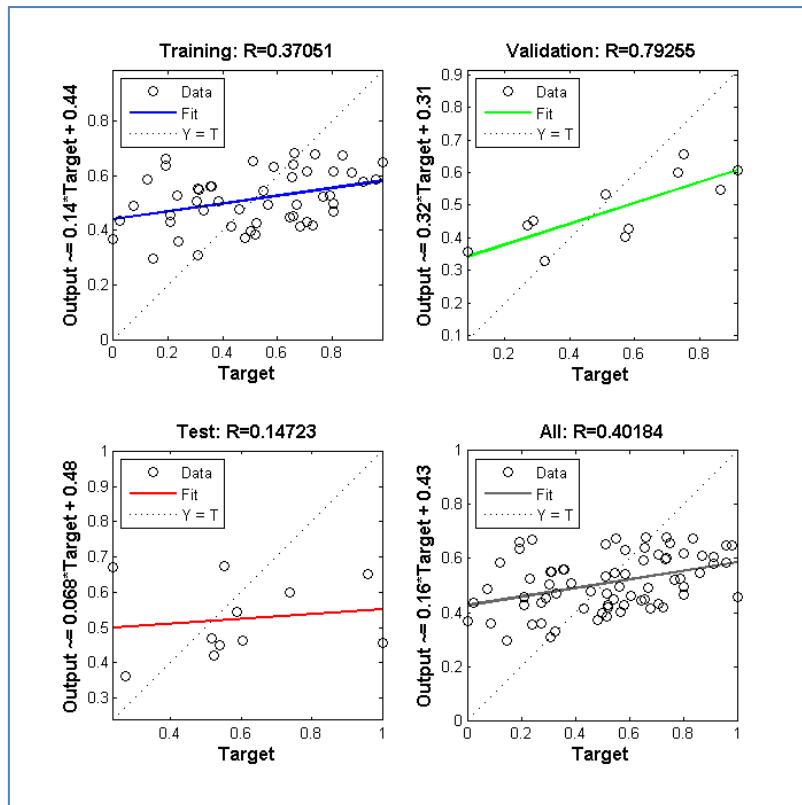
Στα σχήματα 41 - 43 εμφανίζεται η διασπορά των προβλέψεων του νευρωνικού δικτύου σε σχέση με τα πειραματικά δεδομένα για τους εναλλακτικούς αλγορίθμους εκπαίδευσης.



Σχήμα 41: Διασπορά προβλέψεων νευρωνικού δικτύου μίας κρυμμένης στοιβάδας και 6 νευρώνων με τον αλγόριθμο CGB (βλ. κεφάλαιο 4)



Σχήμα 42: Διασπορά προβλέψεων νευρωνικού δικτύου μίας κρυμμένης στοιβάδας και 6 νευρώνων με τον αλγόριθμο CGF (βλ. κεφάλαιο 4)



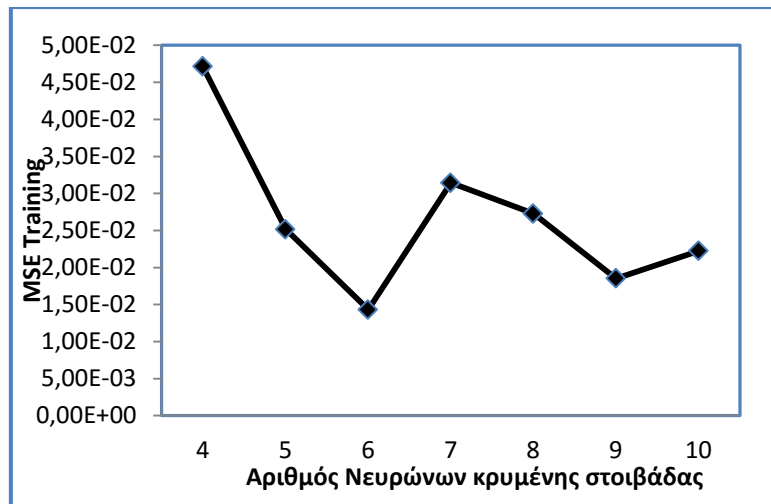
Σχήμα 43: Διασπορά προβλέψεων νευρωνικού δικτύου μίας κρυμμένης στοιβάδας και 6 νευρώνων με τον αλγόριθμο CGP (βλ. κεφάλαιο 4)

Σε όλες τις δοκιμές που πραγματοποιήθηκαν, ο αλγόριθμος Levenberg-Marquardt φάνηκε να δίνει καλύτερα αποτελέσματα.

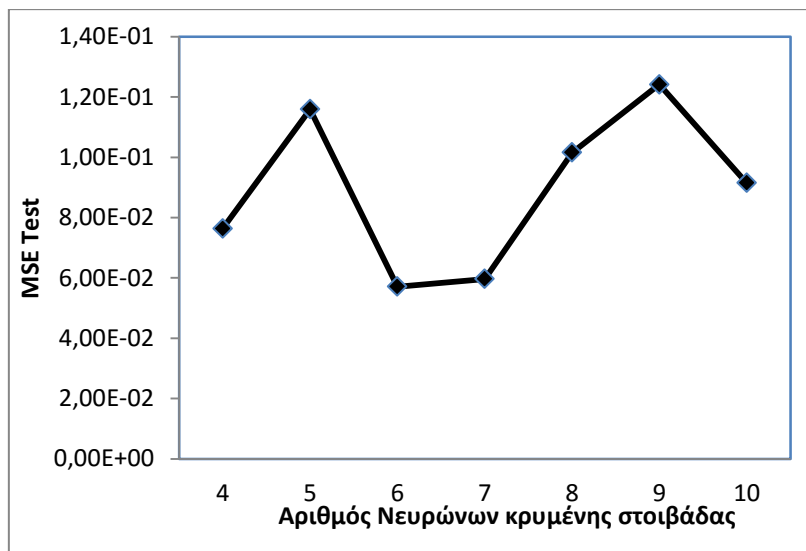
### 5.1.2 Επίδραση Αρχιτεκτονικής Νευρωνικού δικτύου

Ένας από τους σημαντικότερους παράγοντες στην επίδοση ενός νευρωνικού δικτύου είναι η αρχιτεκτονική του. Στα σχήματα 44, 45 φαίνεται το σφάλμα πρόβλεψης (MSE) νευρωνικών δικτύων πρόσθιας τροφοδότησης μίας κρυμμένης στοιβάδας για τα δεδομένα training και για τα δεδομένα testing.

Στα δεδομένα training (σχήμα 44) είναι εμφανές ότι με την αύξηση του αριθμού των νευρώνων, το σφάλμα έχει την τάση να μειωθεί. Αυτό οδηγεί σε overfitting, το οποίο φαίνεται στην αύξηση σφάλματος στα δεδομένα testing (σχήμα 45).

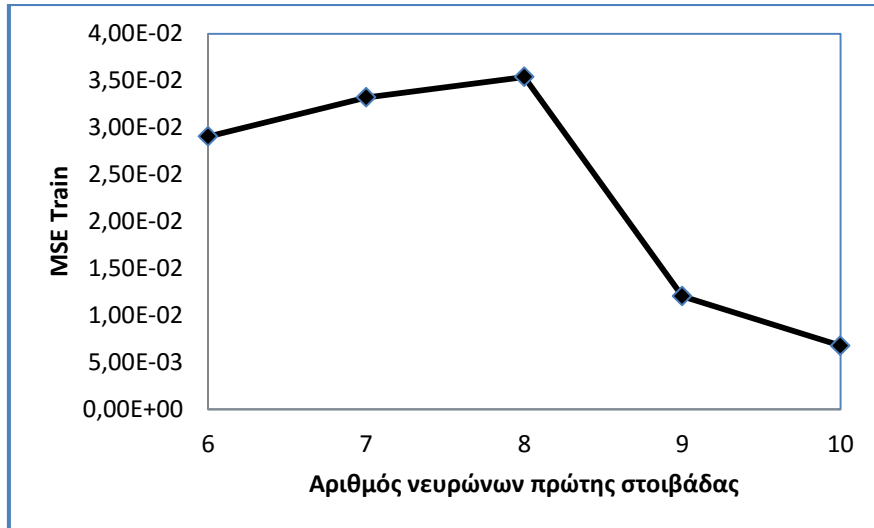


Σχήμα 44: Σφάλμα MSE για τα δεδομένα training για νευρωνικά δίκτυα πρόσθιας τροφοδότησης μίας κρυμμένης στοιβάδας (δοκιμή 1)

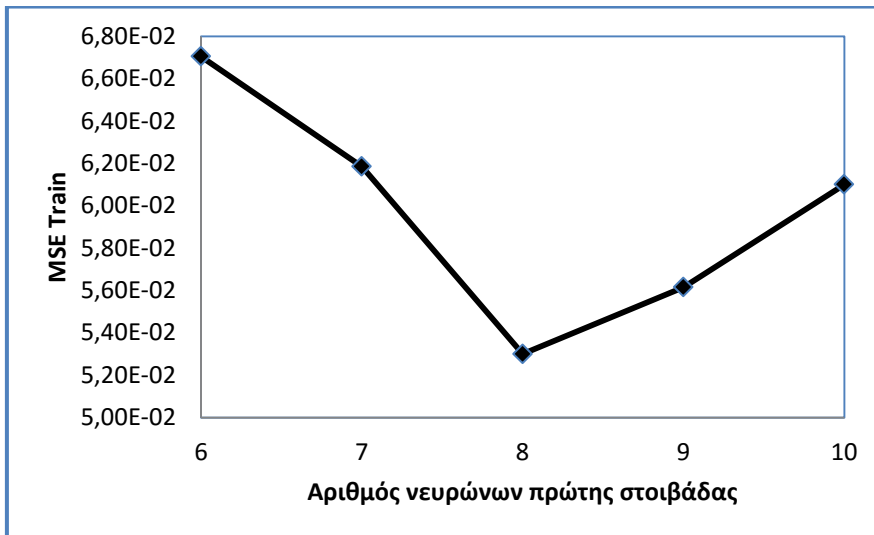


Σχήμα 45: Σφάλμα MSE για τα δεδομένα testing για νευρωνικά δίκτυα πρόσθιας τροφοδότησης μίας κρυμμένης στοιβάδας (δοκιμή 1)

Αντίστοιχα διαγράμματα για νευρωνικά δίκτυα πρόσθιας τροφοδότησης δύο κρυμμένων στοιβάδων νευρώνων, όπου η δεύτερη στοιβάδα έχει δύο νευρώνες λιγότερους από την πρώτη απεικονίζονται στα σχήματα 46, 47.



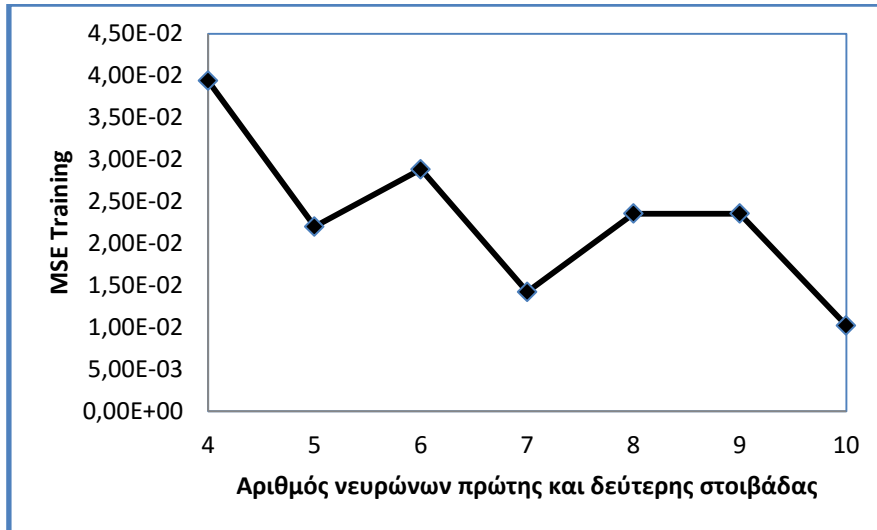
**Σχήμα 46:** Σφάλμα MSE για τα δεδομένα training για νευρωνικά δίκτυα πρόσθιας τροφοδότησης δύο κρυμμένων στοιβάδων όπου ο αριθμός νευρώνων της πρώτης στοιβάδας είναι κατά δύο μεγαλύτερος από αυτόν της δεύτερης (δοκιμή 1)



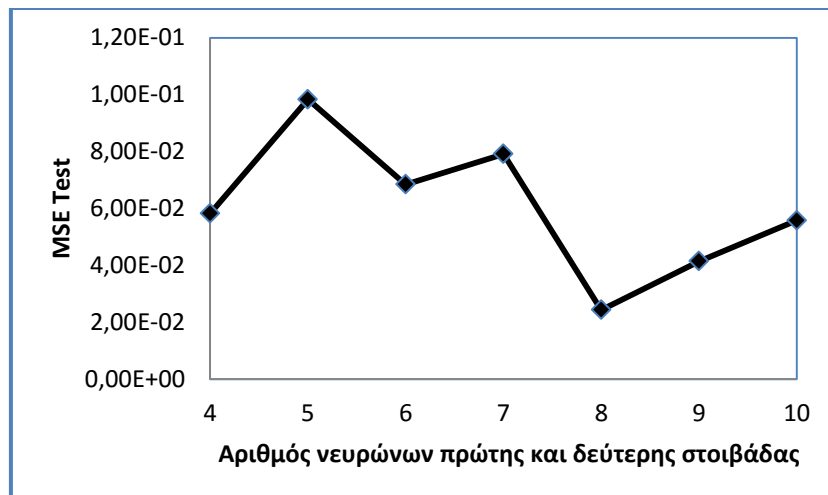
**Σχήμα 47:** Σφάλμα MSE για τα δεδομένα testing για νευρωνικά δίκτυα πρόσθιας τροφοδότησης δύο κρυμμένων στοιβάδων όπου ο αριθμός νευρώνων της πρώτης στοιβάδας είναι κατά δύο μεγαλύτερος από αυτόν της δεύτερης (δοκιμή 1)

Στα διαγράμματα 48, 49 απεικονίζεται το MSE για νευρωνικά δίκτυα πρόσθιας τροφοδότησης όπου ο αριθμός νευρώνων της δεύτερης στοιβάδας είναι ίσος με τον αριθμό της πρώτης.





Σχήμα 48: Σφάλμα MSE για τα δεδομένα training για νευρωνικά δίκτυα πρόσθιας τροφοδότησης δύο κρυμμένων στοιβάδων όπου ο αριθμός νευρώνων της πρώτης στοιβάδας είναι ίσος με αυτόν της δεύτερης (δοκιμή 1)



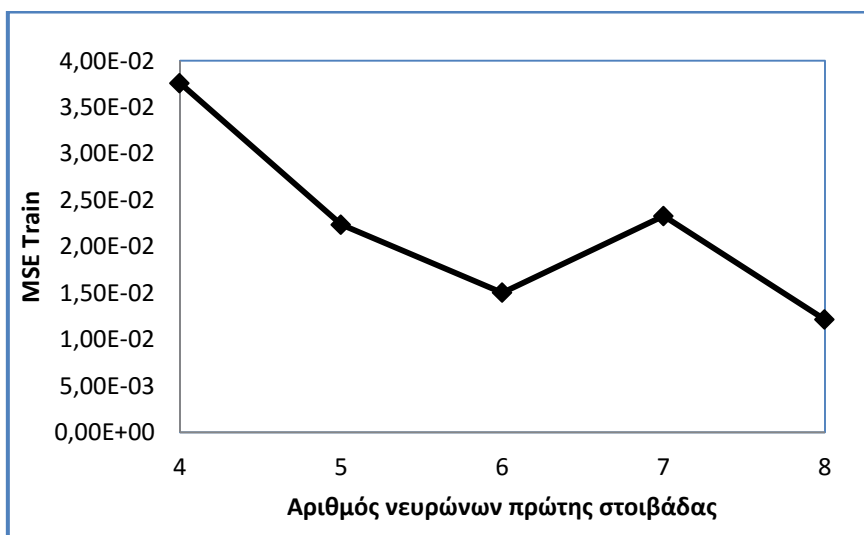
Σχήμα 49: Σφάλμα MSE για τα δεδομένα testing για νευρωνικά δίκτυα πρόσθιας τροφοδότησης δύο κρυμμένων στοιβάδων όπου ο αριθμός νευρώνων της πρώτης στοιβάδας είναι ίσος με αυτόν της δεύτερης (δοκιμή 1)

Τέλος, στα διαγράμματα 50, 51 απεικονίζεται το MSE για νευρωνικά δίκτυα πρόσθιας τροφοδότησης όπου ο αριθμός νευρώνων της δεύτερης στοιβάδας είναι μεγαλύτερος κατά δύο από τον αριθμό της πρώτης.

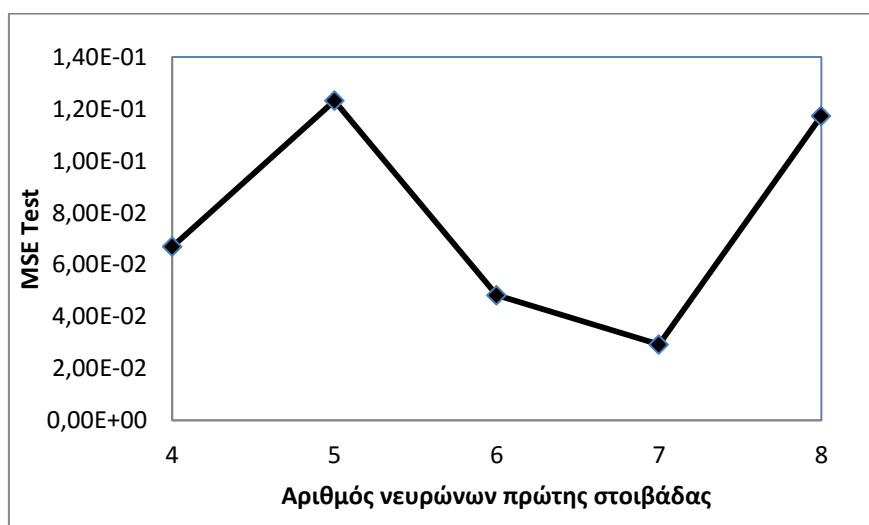
Το βέλτιστο δίκτυο, επελέγη με βάση το συνολικό σφάλμα:

$$MSE_{total} = MSE_{train} + MSE_{validation} + MSE_{test}$$

Όπως φαίνεται από τους πίνακες 11-15, το βέλτιστο νευρωνικό δίκτυο πρόσθιας τροφοδότησης με βάση αυτό το κριτήριο είναι το νευρωνικό δίκτυο μίας κρυμμένης στοιβάδας, με 6 νευρώνες σε αυτήν κατά την δοκιμή 1.



**Σχήμα 50:** Σφάλμα MSE για τα δεδομένα training για νευρωνικά δίκτυα πρόσθιας τροφοδότησης δύο κρυμμένων στοιβάδων όπου ο αριθμός νευρώνων της πρώτης στοιβάδας είναι κατά δύο μικρότερος από αυτόν της δεύτερης (δοκιμή 1)



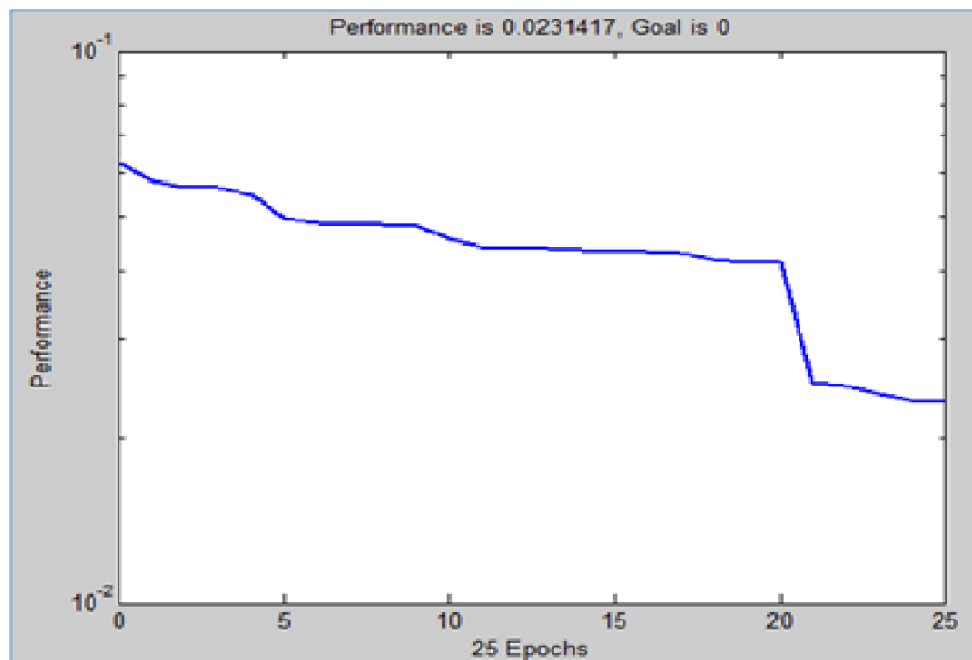
**Σχήμα 51:** Σφάλμα MSE για τα δεδομένα testing για νευρωνικά δίκτυα πρόσθιας τροφοδότησης δύο κρυμμένων στοιβάδων όπου ο αριθμός νευρώνων της πρώτης στοιβάδας είναι κατά δύο μικρότερος από αυτόν της δεύτερης (δοκιμή 1)

Από τα παραπάνω διαγράμματα, είναι εμφανής η τάση του σφάλματος να μειώνεται στα δεδομένα training με την αύξηση του αριθμού νευρώνων. Υπερβολική ελάττωση του σφάλματος στα δεδομένα του training, οδηγεί σε υπέρ-προσαρμογή

των δεδομένων (overfitting) και αυτό απεικονίζεται σε τάση αύξησης του σφάλματος στα δεδομένα testing.

## 5.2 Αποτελέσματα νευρωνικών δικτύων ακτινικής βάσης

Στο σχήμα 52, εμφανίζεται η ελάττωση του σφάλματος στα δεδομένα training κατά την εκπαίδευση νευρωνικού δικτύου ακτινικής βάσης.



Σχήμα 52: Σφάλμα δεδομένων training κατά την εκπαίδευση νευρωνικών δικτύων ακτινικής βάσης

Αριθμός νευρώνων	Spread	MSE Train	MSE test	R train	R test
5	0.250	0.0554	0.0671	0.1171	0.1353
5	0.500	0.0492	0.0578	0.2148	0.1809
5	1.000	0.0533	0.0837	0.1499	0.0048
5	1.500	0.0492	0.1434	0.2156	0.1065
5	2.000	0.0499	0.1460	0.2050	0.0950
10	0.250	0.0507	0.0571	0.1908	0.1332
10	0.500	0.0448	0.0569	0.2848	0.1115
10	1.000	0.0466	0.1531	0.2565	0.2081
10	1.500	0.0473	0.1389	0.2454	0.2900
10	2.000	0.0455	0.3156	0.2736	0.2667
15	0.250	0.0437	0.0463	0.3023	0.3108
15	0.500	0.0407	0.4972	0.3514	0.0023
15	1.000	0.0422	0.4605	0.3276	0.4227
15	1.500	0.0440	1.3400	0.2990	0.6657
15	2.000	0.0433	0.3977	0.3098	0.8240
20	0.250	0.0373	0.0715	0.4049	0.1637
20	0.500	0.0245	2.6981	0.6087	0.0261
20	1.000	0.0236	1.3732	0.6231	0.2833
20	1.500	0.0389	4.8283	0.3790	0.8163
20	2.000	0.0412	0.9692	0.3429	0.6982
25	0.250	0.0317	0.0684	0.4940	0.2584
25	0.500	0.0229	2.5625	0.6348	0.0302
25	1.000	0.0195	1.7455	0.6896	0.4163
25	1.500	0.0376	5.6447	0.4001	0.8029
25	2.000	0.0231	1.9188	0.6309	0.3693
30	0.250	0.0217	0.0993	0.7126	0.5090
30	0.500	0.0124	0.0590	0.8413	0.7011
30	1.000	0.0138	0.0668	0.8170	0.6817
30	2.000	0.0219	0.1054	0.6912	0.4937
40	0.250	0.0108	0.0518	0.8622	0.7185
40	0.500	0.0105	0.0503	0.8663	0.7219
40	1.000	0.0088	0.0387	0.8893	0.7411
40	2.000	0.0106	0.0506	0.8655	0.7213
50	0.250	0.0033	0.0131	0.9602	0.8802
50	0.500	0.0036	0.0142	0.9588	0.8768
50	1.000	0.0069	0.0304	0.9141	0.8486
50	2.000	0.0095	0.0421	0.8790	0.7325
60	0.250	0.0021	0.0084	0.9747	0.8942
60	0.500	0.0022	0.0088	0.9735	0.8932
60	1.000	0.0020	0.0081	0.9756	0.8960
60	2.000	0.0047	0.0187	0.9426	0.8648

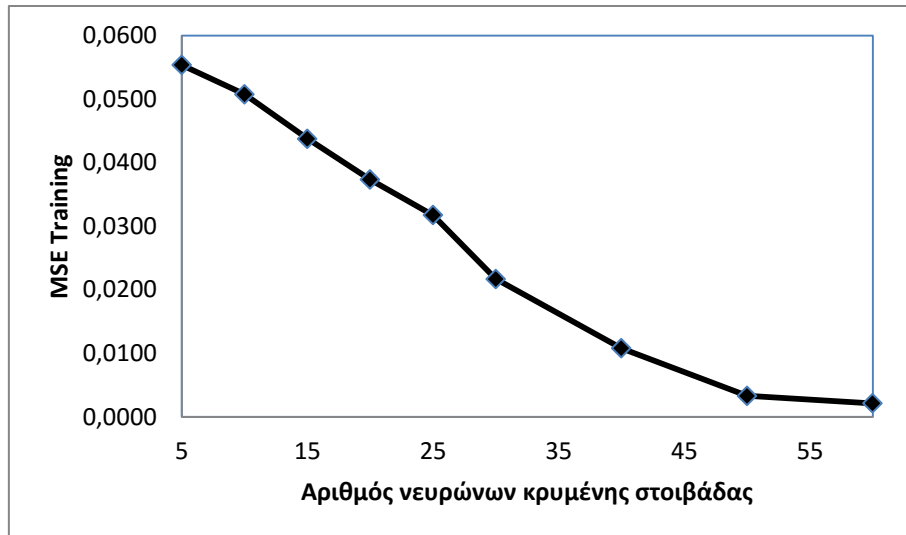
Πίνακας 17: Δοκιμές νευρωνικών δικτύων ακτινικής βάσης

### 5.2.1 Επίδραση Αρχιτεκτονικής νευρωνικών δικτύων ακτινικής βάσης

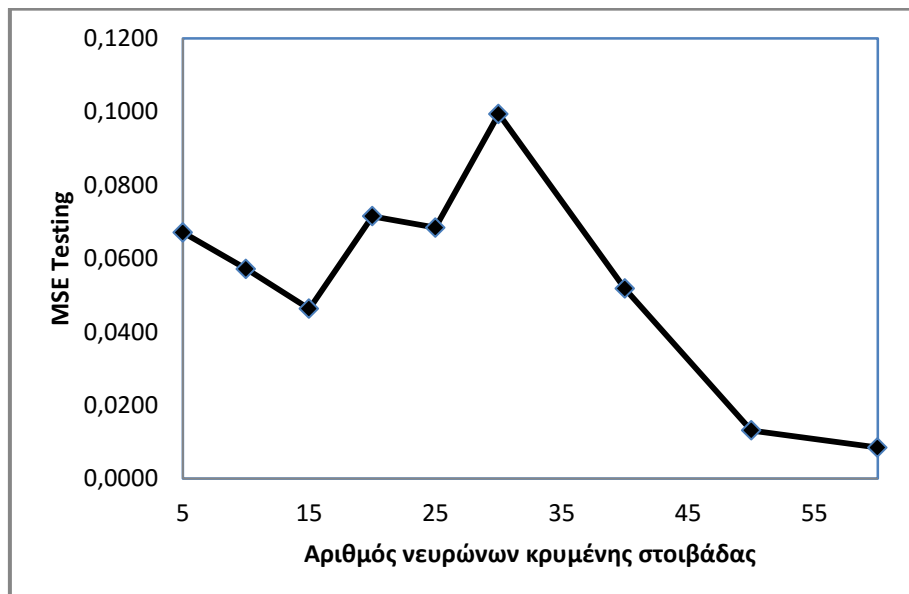
Όπως αναφέρθηκε σε προηγούμενο κεφάλαιο, για τα νευρωνικά δίκτυα ακτινικής βάσης έγινε μία δοκιμή σε κάθε αρχιτεκτονική. Κάθε αρχιτεκτονική νευρωνικού δικτύου ακτινικής βάσης, δοκιμάστηκε για διάφορες τιμές spread της συνάρτησης ενεργοποίησης (συνάρτηση Gauss). Τελικά επελέγη η καλύτερη

αρχιτεκτονική (αριθμός νευρώνων, spread) με βάση το μικρότερο σφάλμα και τον συντελεστή συσχέτισης R πλησιέστερα στην μονάδα.

Στο σχήμα 53, παρατίθεται η επίδραση του αριθμού νευρώνων, για spread 0.25 στο σφάλμα δεδομένων εκπαίδευσης (MSE training), ενώ στο σχήμα 54 η επίδραση του αριθμού νευρώνων στο σφάλμα δεδομένων testing.



Σχήμα 53: Επίδραση αριθμού νευρώνων στο σφάλμα δεδομένων training

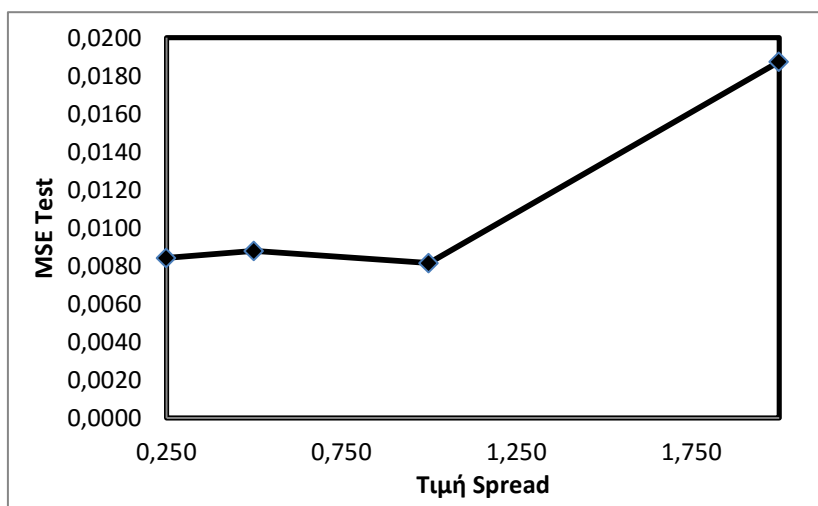


Σχήμα 54: Επίδραση αριθμού νευρώνων στο σφάλμα δεδομένων testing

Όπως και στα νευρωνικά δίκτυα πρόσθιας τροφοδότησης, η αύξηση του αριθμού νευρώνων τείνει να ελαττώνει το σφάλμα εκπαίδευσης. Παρ' όλο που αναμενόταν αύξηση του σφάλματος στα δεδομένα testing μετά από έναν αριθμό νευρώνων,

βρέθηκε ότι το δίκτυο μέγιστου αριθμού νευρώνων κρυμμένης στοιβάδας (60 νευρώνες) που δοκιμάστηκε είχε το μικρότερο σφάλμα.

Όπως αναφέρθηκε, για τα νευρωνικά δίκτυα ακτινικής βάσης, δοκιμάστηκαν διαφορετικές τιμές spread. Ενδεικτικά παρουσιάζεται η επίδραση του spread στο σφάλμα testing (MSE Testing) για το βέλτιστο δίκτυο στο σχήμα 55.



Σχήμα 55: Επίδραση spread στο MSE Test για το βέλτιστο νευρωνικό δίκτυο

## 6. Συμπεράσματα/Προτάσεις για μελλοντική έρευνα

Από την ανάλυση που παρουσιάστηκε στα προηγούμενα κεφάλαια, προκύπτουν τα εξής συμπεράσματα:

- Με χρήση νευρωνικών δικτύων (πρόσθιας τροφοδότησης και ακτινικής βάσης), ήταν εφικτή η κατασκευή στατιστικών μαθηματικών μοντέλων για την κατεργασία της λείανσης μετάλλων. Τα στατιστικά μοντέλα που προέκυψαν, και στα δύο είδη νευρωνικών δικτύων, ήταν σε θέση να δώσουν ικανοποιητική ακρίβεια πρόβλεψης για δεδομένα όπου η μαθηματική μοντελοποίηση με ντετερμινιστικές μεθόδους θα ήταν ιδιαίτερα δυσχερής.
- Με βάση τις δοκιμές που έγιναν, προκύπτει ότι με χρήση νευρωνικών δικτύων ακτινικής βάσης, είναι δυνατόν να προκύψουν μαθηματικά μοντέλα της κατεργασίας λείανσης, τα οποία παρουσιάζουν σημαντικά χαμηλότερα σφάλματα (βλ. πίνακες 11 έως 17) σε σχέση με τα νευρωνικά δίκτυα πρόσθιας τροφοδότησης.

- Λόγω της απλότητας της αρχιτεκτονικής των νευρωνικών δικτύων ακτινικής βάσης, έχουν μικρότερο υπολογιστικό κόστος από τα νευρωνικά δίκτυα πρόσθιας τροφοδότησης. Μειωμένο υπολογιστικό κόστος καθιστά ευχερέστερο έναν μεγαλύτερο αριθμό δοκιμών, ώστε να βρεθεί το βέλτιστο νευρωνικό δίκτυο ακτινικής βάσης για το εκάστοτε πρόβλημα. Επίσης, τα καθιστά ένα ισχυρότερο εργαλείο, όταν προκύπτουν προβλήματα εξαιρετικά αυξημένου υπολογιστικού κόστους.

Πιθανές προεκτάσεις τις παρούσας εργασίας θα ήταν δυνατόν να έχουν τις εξής κατευθύνσεις:

- Καθώς τα νευρωνικά δίκτυα ακτινικής βάσης, καθώς επίσης και τα νευρωνικά δίκτυα πρόσθιας τροφοδότησης είναι στατιστικά μοντέλα, επόμενο βήμα θα ήταν δυνατόν να είναι η επέκταση της διαδικασίας που ακολουθήθηκε, για μεγαλύτερο αριθμό πειραματικών δεδομένων
- Οι δοκιμές πραγματοποιήθηκαν για τρία είδη χάλυβα και έξι είδη λειαντικών τροχών. Επόμενη προέκταση είναι η επέκταση του στατιστικού μοντέλου, ώστε να περιλαμβάνει περισσότερα είδη λειαντικών τροχών και χάλυβα.
- Τέλος, λόγω του γεγονότος ότι τα νευρωνικά δίκτυα είναι στατιστικό μοντέλο, χωρίς να υπεισέρχεται στην φυσική του προβλήματος, οι τεχνικές αυτές είναι δυνατόν να χρησιμοποιηθούν και για μαθηματική μοντελοποίηση άλλων κατεργασιών όπου η ντετερμινιστική προσέγγιση είναι δυσχερής.

## Βιβλιογραφία

- [1] Valery Marinov, Manufacturing Technology, Abir Roy, Aug. 27 2011
- [2] Kalpakjian Schmid, Manufacturing Processes for Engineering Materials, 5th ed., 2008 Pearson Education
- [3] Σημειώσεις Μαθήματος Κατεργασίες ΙΙ, Σχολής Μηχανολόγων Μηχανικών Ε.Μ.Π. <http://courseware.mech.ntua.gr/ml26076/mathimata/2006Liansi.pdf>
- [4] Γεώργιος Χρυσολουρής, Βασικές Αρχές Λείανσης, Πανεπιστήμιο Πατρών, Τμήμα Μηχανολόγων & Αεροναυπηγών Μηχανικών, Σημειώσεις, Πάτρα 2003, <https://eclass.upatras.gr/modules/document/file.php/MECH1111/%CE%95%CE%BA%CF%80%CE%B1%CE%B9%CE%B4%CE%B5%CF%85%CF%84%CE%B9%CE%BA%CF%8C%20%CE%A5%CE%BB%CE%B9%CE%BA%CF%8C/Basics%20of%20Grinding.pdf>
- [5] Raul Rojas, Neural Networks, A Systematic Introduction, Springer Verlag Berlin 1996
- [6] David Kriesel, A brief Introduction to Neural Networks, Pearson, Prentice Hall
- [7] Simon Haykin, McMaster University, Hamilton, Ontario, Canada, Neural Networks and Learning Machines, 2nd Edition 1999
- [8] Τζίντζας Χρήστος, Μοντελοποίηση των δυνάμεων κοπής στην κατεργασία του τορναρίσματος με την χρήση μοντέλων νευρωνικών δικτύων, Διπλωματική Εργασία, Σχολή Μηχανολόγων Μηχανικών Ε.Μ.Π., Αθήνα 2003
- [9] Ηρώα Μαρία, Ανάπτυξη μοντέλων Νευρωνικών Δικτύων για την προσομοίωση της διαμόρφωσης κυλινδρικών δοκιμίων υπό την επίδραση θλιπτικών αξονικών δυνάμεων, Διπλωματική Εργασία, Σχολή Μηχανολόγων Μηχανικών Ε.Μ.Π., Αθήνα 2003
- [10] G. P. J Shmitz, C. Aldrich, “Neurofuzzy modeling of chemical process systems with ellipsoidal radial basis functions neural networks and genetic algorithms”, Computers and Chemical Engineering Vol. 22, pp. S1001-1004, (1998)



- [11] James A. Leonard, Mark A. Kramer, „Radial Basis Function Networks for Classifying Process Faults”, American Control Conference, San Diego, CA, May (1990)
- [12] S. Chen, C. F. N. Cowan, “Orthogonal Least Squares Learning Algorithm for Radial Basis Neural Networks” IEEE Transactions on Neural Networks, vol. 2, No 2, March 1991
- [13] Christian Darken and John Moody, “Fast Adaptive K-Means Clustering: Some Empirical Results”, IEEE INNS International J. Conference on Neural Networks, San Diego (1990)
- [14] Markopoulos AP. Simulation of grinding by means of the Finite Element Method and Artificial Neural Networks. Computational Methods for Optimizing Manufacturing Technology (ed. Davim JP), IGI Global Hershey, PA, USA. 2011: 193-218. doi: 10.4018/978-1-4666-0128-4.ch008