ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΑΡΑΓΩΓΗΣ ΚΑΙ ΔΙΟΙΚΗΣΗΣ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΠΡΟΒΛΕΨΗ ΤΙΜΩΝ ΤΡΟΦΙΜΩΝ ΜΕ ΧΡΗΣΗ ΤΗΣ ΕΜΠΕΙΡΙΚΗΣ ΜΕΘΟΔΟΥ ΑΠΟΣΥΝΘΕΣΗΣ (EMD) ΚΑΙ ΤΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ

ΒΙΚΤΩΡΙΑ ΣΠΗΛΙΑΔΗ

ΧΑΝΙΑ 2011

Ευχαριστίες

Σε αυτό το σημείο, θα ήθελα να ευχαριστήσω τον κ. Γεώργιο Ατσαλάκη, διδάκτορα του Πολυτεχνείου Κρήτης για την συμπαράσταση και τη βοήθειά του κατά τη διάρκεια εκπόνησης της παρούσας διπλωματικής εργασίας.

Επίσης θέλω να ευχαριστήσω τους φίλους μου, και ιδιαίτερα τον Ευτύχη Πρωτοπαπαδάκη για την πολύτιμη βοήθειά του.

Τέλος, θέλω να ευχαριστήσω θερμά την οικογένειά μου, που όλα αυτά τα χρόνια με υποστήριζε και μου έδινε δύναμη.

ΠΕΡΙΕΧΟΜΕΝΑ

ΚΕΦΑΛΑΙΟ 1: Εισαγωγή

Στην εμπορική πραγματικότητα που ζούμε τα τελευταία χρόνια, η κατανάλωση μειώνεται όλο και περισσότερο. Αυτό επηρεάζει άμεσα τα κέρδη των επιχειρήσεων. Ανάλογα με την εποχή και τις συνθήκες που επικρατούν, οι τιμές για κάθε τι γύρω μας μεταβάλλονται συνεχώς. Αυτές οι μεταβολές όμως επηρεάζουν πολύ την λειτουργία των επιχειρήσεων. Μια επιχείρηση για να πετύχει μεγάλο κέρδος, πρέπει να μπορεί να διαχειριστεί σωστά αυτές τις μεταβολές. Αν μπορούσε να γνωρίζει πως πρόκειται να κυμανθούν οι τιμές, θα μπορούσε να παίρνει πιο κερδοφόρες αποφάσεις όσον αφορά την παραγωγή-εισαγωγή προϊόντων, την αποθεματοποίηση, την προώθηση-πώληση προϊόντων κ.α.. Θα μπορούσε να συμβαδίζει τις ανάγκες της και τους στόχους της ανάλογα με τις οικονομικές εξελίξεις ώστε να έχει μεγαλύτερη απόδοση.

Είναι προφανές λοιπόν, ότι η πρόβλεψη των τιμών αποτελεί σημαντικό παράγοντα επιτυχίας στον χώρο των επιχειρήσεων και όχι μόνο. Για αυτό τον λόγο η έρευνα γύρω από μοντέλα πρόβλεψης είναι κάτι απασχολεί τους ερευνητές όλο και περισσότερο τα τελευταία χρόνια.

Ένα μοντέλο πρόβλεψης που έχει αναπτυχθεί και εφαρμόζεται ευρέως, είναι τα νευρωνικά δίκτυα. Οι τιμές των προϊόντων εξαρτώνται από πολλούς παράγοντες που δεν μπορούν να προβλεφθούν, όπως κοινωνικούς και πολιτικούς παράγοντες. Τα νευρωνικά δίκτυα ξεχωρίζουν για αυτήν τους την ιδιότητα, να επεξεργάζονται μη γραμμικά και μη στατικά σήματα.

Στην παρούσα εργασία προτείνεται το μοντέλο EMD-BPN για την πρόβλεψη των τιμών κάποιων προϊόντων. Ένα μοντέλο δηλαδή που συνδυάζει τα νευρωνικά δίκτυα με την *εμπειρική μέθοδο αποσύνθεσης* (EMD). Αυτό το μοντέλο πρόβλεψης πρώτα χρησιμοποιεί την EMD για να αποσυνθέσει τα περίπλοκα αρχικά δεδομένα σε ένα πεπερασμένο σύνολο εσωτερικών λειτουργιών (IMFs) με απλούστερη συχνότητα, και σε κάποιο υπόλοιπο, και έπειτα χρησιμοποιεί τα IMFs και το υπόλοιπο, ως εισόδους για τα νευρωνικά δίκτυα στα οποία γίνεται η μοντελοποίηση και η πρόβλεψη των αποτελεσμάτων. Η τελική τιμή πρόβλεψης είναι το άθροισμα όλων αυτών των αποτελεσμάτων.

Θα προσπαθήσουμε λοιπόν να δούμε πως ανταποκρίνεται η μέθοδος στο συγκεκριμένο πρόβλημα. Για να το ελέγξουμε αυτό, θα θεωρήσουμε ένα μέρος των δεδομένων ως δείγμα εκπαίδευσης και το υπόλοιπο σύνολο δεδομένων ως το δείγμα δοκιμής. Θα εκπαιδεύσουμε δηλαδή το μοντέλο χρησιμοποιώντας τα πρώτα στοιχεία, και έπειτα θα συγκρίνουμε τις τιμές πρόβλεψης που θα δώσει το μοντέλο με τις πραγματικές τιμές των δεδομένων.

ΚΕΦΑΛΑΙΟ 2: Εμπειρική μέθοδος αποσύνθεσης (Empirical mode decomposition-EMD)

2.1. Εισαγωγή

Η εμπειρική μέθοδος αποσύνθεσης δημοσιεύθηκε το 1996, στη μελέτη των Norden Huang et al., που αφορούσε μια νέα μέθοδο ανάλυσης σημάτων. Η βασική αρχή της μεθόδου είναι ότι απλοποιεί μια χρονοσειρά, σε ένα άθροισμα συναρτήσεων διακύμανσης.

Σύμφωνα με τη μελέτη, η μέθοδος αυτή, δεν χρησιμοποιεί κάποια προηγούμενη βάση προσέγγισης, όπως π.χ. στον μετασχηματισμό Fourier, αλλά είναι καθοδηγούμενη από το ίδιο το σήμα. Αυτή είναι και η κύρια διαφορά της από τις άλλες μεθόδους αποσύνθεσης. Έτσι, η μέθοδος αυτή μπορεί να εφαρμοστεί σε οποιαδήποτε δεδομένα, ακόμα κι αν αυτά προκύπτουν από μη στάσιμες ή μη γραμμικές διαδικασίες, με πλήρη φυσική σημασία των αποτελεσμάτων.

Όταν η μέθοδος EMD συνδυάζεται με την παραδοσιακή Hilbert spectral analysis (ανάλυση φάσματος με χρήση μετασχηματισμού Hilbert), ονομάζεται Hilbert–Huang Transform (HHT). Η χρήση του μετασχηματισμού Hilbert επιτρέπει τον υπολογισμό στιγμιαίων συχνοτήτων και πλάτους ως συναρτήσεις χρόνου για κάθε εξαγόμενο περιοδικό συστατικό.

Οποιοδήποτε σήμα περιέχει ένα σύνολο εγγενών τάσεων ταλάντωσης. Η μέθοδος αυτή, βασισμένη στη αρχή του διαχωρισμού κλιμάκων, μπορεί να προσδιορίζει αυτές τις τάσεις και να τις εξάγει. Επιπλέον, μπορεί να εξάγει και ταλαντώσεις που συνυπάρχουν ταυτόχρονα.

Κάθε μια από τις παραπάνω εγγενείς τάσεις ταλάντωσης, αναπαρίσταται από μία συνάρτηση , την λεγόμενη Intrinsic Mode Function (IMF).

Τα IMFs πρέπει να πληρούν δύο προϋποθέσεις:

1)Ο αριθμός των ακροτάτων και ο αριθμός των σημείων που περνούν από το μηδέν πρέπει να διαφέρουν κατά 1, και
2) ο τοπικός μέσος να είναι μηδέν (δηλαδή ο μέσος όρος της επάνω και της κάτω καμπύλης να είναι μηδέν).

Η IMF είναι το αντίστοιχο της αρμονικής συνάρτησης για τις άλλες μεθόδους αποσύνθεσης, με τη σημαντική διαφορά ότι και το πλάτος και η συχνότητα είναι συναρτήσεις του χρόνου.

Ο αλγόριθμος της EMD παρουσιάζεται ως εξής:

Βήμα-1:

Προσδιορίζουμε τα τοπικά ακρότατα των πειραματικών δεδομένων. $\{x(t)\}$

Συνδέουμε όλα τα τοπικά μέγιστα με μια κυβική Spline, η οποία αποτελεί την ανώτερη καμπύλη των δεδομένων. $\{U(t)\}$

Επαναλαμβάνουμε την ίδια διαδικασία για τα τοπικά ελάχιστα και υπολογίζουμε την κατώτερη καμπύλη των δεδομένων μας. $\{L(t)\}$

Όλα τα δεδομένα μας πρέπει τώρα να περικλείονται από τις δύο καμπύλες.

Ο μέσος όρος της επάνω και κάτω καμπύλης συμβολίζεται $m_1(t)$ και υπολογίζεται από τον τύπο:

$m_1(t)=U(t)+L(t)/2$ **(1)**

Αφαιρούμε την τρέχουσα μέση τιμή $m_1(t)$ από την αρχική χρονοσειρά x(t) και παίρνουμε το πρώτο συστατικό $h_1(t)$. Δηλαδή,

$h_1(t)=x(t)-m_1(t)$ **(2)**

Αν το στοιχείο $h_1(t)$ έχει όλα τα μέγιστα θετικά και όλα τα ελάχιστα αρνητικά, τότε είναι IMF και το βήμα-1 σταματάει εδώ. (Επίσης θα μπορούσαμε να σταματήσουμε εδώ αν θέλαμε να βρούμε κυματομορφές με συγκεκριμένο εύρος διακοπής για φυσική μελέτη).

Αν το στοιχείο $h_1(t)$ δεν είναι IMF η διαδικασία του κοσκινίσματος πρέπει να επαναληφθεί όσες φορές χρειάζεται μέχρι τα εξαγόμενα σήματα να αναχθούν σε IMF. Σε αυτή την περίπτωση το $h_1(t)$ αντιμετωπίζεται ως τα νέα δεδομένα για να επαναλάβουμε την παραπάνω διαδικασία.

$h_{11}(t)=h_1(t)- m_{11}(t)$ **(3)**

Αν και πάλι η συνάρτηση $h_{11}(t)$ δεν είναι IMF, η διαδικασία κοσκινίσματος συνεχίζεται για k φορές μέχρι τα επιτρεπόμενα όρια ανοχής:

$h_{1k}(t) = h_{1(k-1)}(t) – m_{1k}(t)$ **(4)**

<u>Βήμα -2:</u>

Εάν οι χρονοσειρές που θα προκύψουν είναι IMF, ορίζουμε

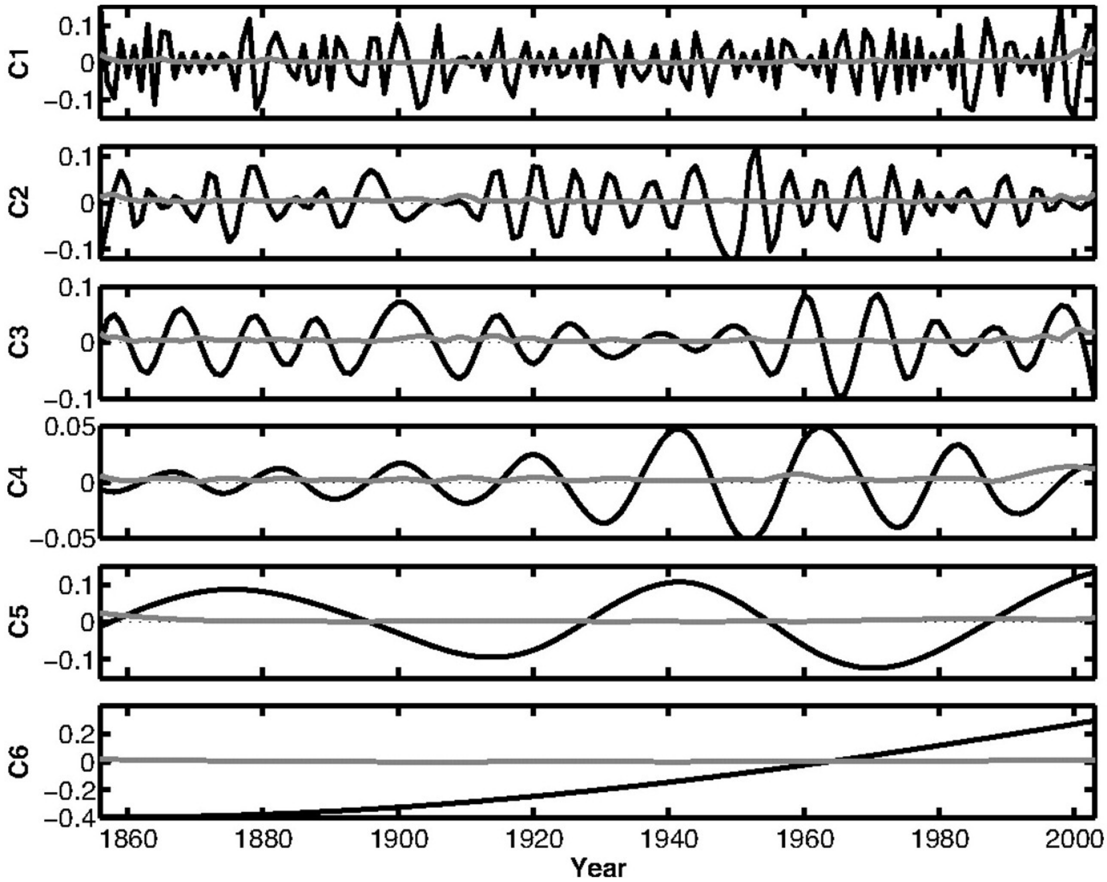
$c_1(t)=h_{1k}(t)$

Έπειτα βρίσκουμε το υπόλειμμα αφαιρώντας το IMF από τα αρχικά δεδομένα. Συμβολίζουμε αυτή τη διαφορά $r_1(t)$.

$r_1(t) =x(t)-c_1(t)$ **(5)**

Το υπόλειμμα $r_1(t)$ λαμβάνεται ως αρχικό δεδομένο και εφαρμόζουμε ξανά το κοσκίνισμα του βήματος 1.

Ακολουθούμε την ίδια διαδικασία μέχρι να βρούμε όλα τα c_i μέχρι και το τελευταίο. Κάθε c_i που θα προκύπτει, θα παρουσιάζει ταλαντώσεις χαμηλότερης συχνότητας από αυτές του προηγούμενου c_{i-1}, (εικόνα 1). Έτσι, το τελικό υπόλειμμα θα είναι μια σταθερή ή μονότονη συνάρτηση που θα αναπαριστά τη γενική τάση της χρονοσειράς.



Εικόνα 1

Τελικά παίρνουμε:

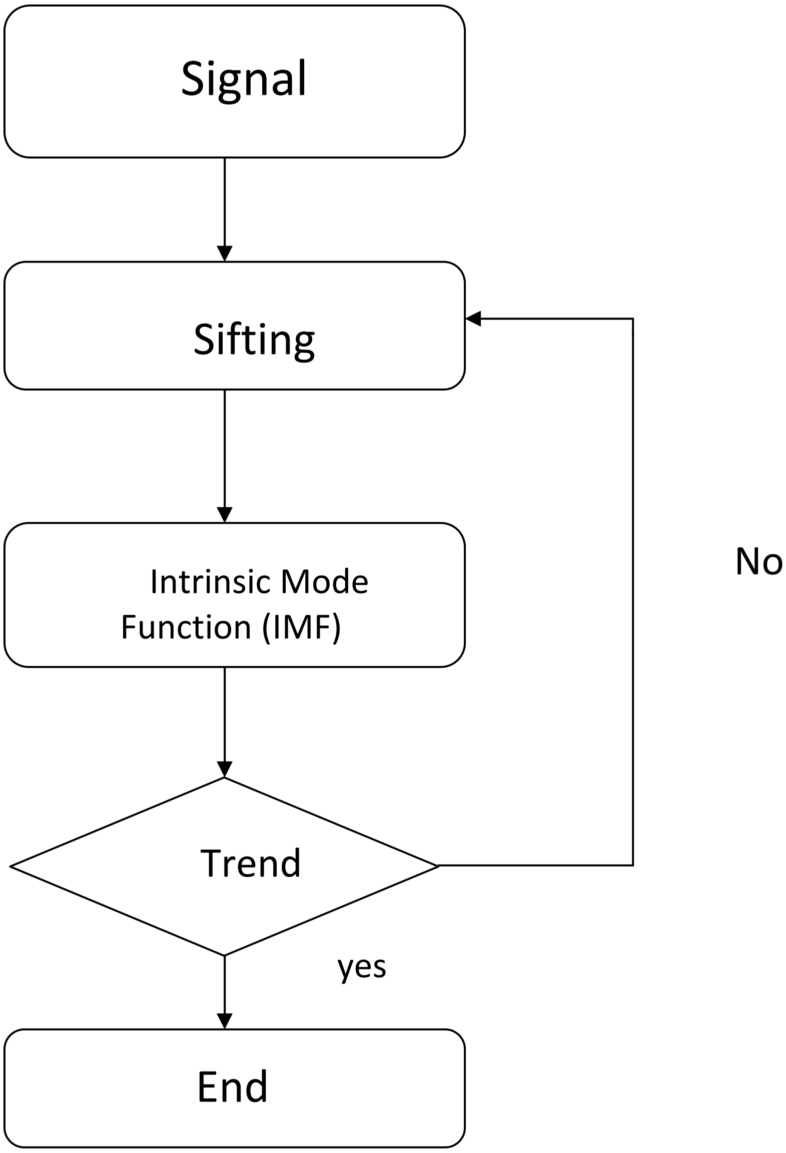
$$x(t) = \sum_{i=1}^{n} c_i\,(t) + r_n\,,$$

$r_i(t) = r_{i-1}(t) - c_i(t)$

Όπου r_n είναι το υπόλειμμα, που αναπαριστά τη μέση τάση του x(t). Όπως είπαμε παραπάνω, τα IMFs $c_1(t)$, $c_2(t)$, $c_3(t)$,……, $c_n(t)$, περιλαμβάνουν διαφορετικές ζώνες συχνότητας που κυμαίνονται από πολύ υψηλές μέχρι πολύ χαμηλές.

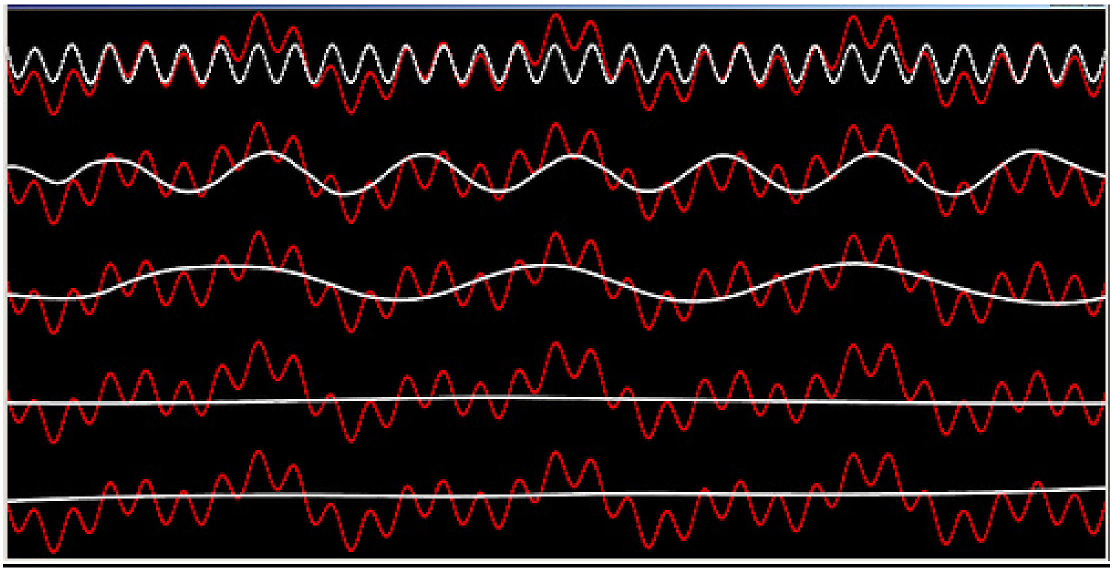
Τα συστατικά συχνότητας που περιέχονται σε κάθε ζώνη συχνότητας είναι διαφορετικά και αλλάζουν με την μεταβολή του σήματος x(t), ενώ η $r_n(t)$ αναπαριστά την κεντρική τάση του x(t).

Παρακάτω ακολουθεί μια σχηματική αναπαράσταση της συνολικής διαδικασίας:



Σχήμα 1

Στην παρακάτω εικόνα βλέπουμε πως εξελίσσονται τα IMFs σε σχέση με το αρχικό σήμα:



Εικόνα 2

2.2. Εφαρμογές EMD

Όπως αναφέραμε, με τη χρήση EMD οποιοδήποτε πολύπλοκο σήμα μπορεί να απλοποιηθεί σε πεπερασμένα IMFs, τα οποία έχουν τμήματα με απλούστερη συχνότητα και δυνατότερη μεταξύ τους αλληλεξάρτηση, πράγμα που τα καθιστά πιο εύκολα και πιο ακριβή στο να προβλεφθούν. Για αυτό το λόγο, το EMD χρησιμοποιείται ευρέως σε πολλά πεδία, όπως για παράδειγμα στην ανάλυση σημάτων των μη στατικών θαλάσσιων κυμάτων σε σεισμικά σήματα και δομικές αναλύσεις, στην κρατική παρακολούθηση γεφυρών και άλλων κατασκευών, στον έλεγχο μηχανημάτων κ.α.

Παρακάτω αναφέρονται μερικά παραδείγματα εφαρμογής της εμπειρικής μεθόδου αποσύνθεσης EMD.

Ευφυής ψηφιακή επεξεργασία σημάτων με μεθόδους EMD, TKO και συνδυασμοί.

Στόχος της εργασίας είναι η εφαρμογή τεχνικών ψηφιακής επεξεργασίας σήματος σε γεωφυσικά σήματα και συγκεκριμένα κατά κύριο λόγο σε σήματα γεωηλεκτρικού δυναμικού μακράς διαρκείας και σε σήματα σεισμικών κυμάτων. Η εργασία εστιάζει στην επεξεργασία της διερεύνησης της εφαρμογής δυο τεχνικών επεξεργασίας ξεχωριστά και σε συνδυασμό. Η πρώτη αφορά την εκτίμηση της στιγμιαίας συχνότητας ενός σήματος μέσω του τελεστή Teager-Kaiser [Maragos_1991] στοχεύοντας σε μια εναλλακτική στα προβλήματα συμβιβασμού της διακριτικής ικανότητας, στην υπολογιστική πολυπλοκότητα, στην ανίχνευση αρνητικών συχνοτήτων και στην εμφάνιση cross-terms άλλων κλασικών μεθόδων. Η δεύτερη που είναι η Empirical Mode Decomposition (EMD) [Huang_1998] αφορά την ανάλυση των δεδομένων σε συνιστώσες AM-FM μορφής καθοδηγούμενη από τα δεδομένα, για την ανάδειξη χαρακτηριστικών δεδομένων μη-στάσιμων και μη-γραμμικών διαδικασιών. Ο συνδυασμός των μεθόδων είναι η εύρεση του συχνοτικού περιεχομένου ενός σήματος με την χρήση του τελεστή Teager- Kaiser από τις επιμέρους συνιστώσες που έχουν προκύψει με την μέθοδο (EMD) και ονομάζεται μετασχηματισμός Teager-Huang και αποτελεί μια εναλλακτική δυνατότητα στην εύρεση του φάσματος ενός σήματος.

Εφαρμογή EMD σε ηλεκτροεγκεφαλογραφήματα (ΗΕΓ)

Σκοπός της εργασίας είναι η μελέτη μεθόδου υπολογισμού της συνδεσιμότητας περιοχών του εγκεφάλου κατά την διάρκεια μιας πραγματικής και μιας φανταστικής κίνησης του δακτύλου. Η μελέτη βασίζεται σε μετρήσεις (καταγραφές) ηλεκτροεγκεφαλογραφημάτων (ΗΕΓ). Η συνδεσιμότητα εκφράζεται ποσοτικά με το μέγεθος της συνάφειας (coherence) και μπορεί να δώσει απάντηση στο κατά πόσο τα σήματα του ηλεκτροεγκεφαλογραφήματος που παίρνουμε από τα δυο κανάλια σχετίζονται μεταξύ τους. Για να επιτευχθεί αυτό υλοποιήθηκε μια μη γραμμική υπολογιστική μέθοδος συγχρονισμού. Η μη γραμμική μέθοδος ανάλυσης και υπολογισμού της συνδεσιμότητας των εγκεφαλικών περιοχών που υλοποιήθηκε έγινε στο μαθηματικό χώρο της φάσης. Σε αυτό το χώρο το σήμα μου είναι ένα δυναμικό σύστημα. Η πληροφορία που περιέχεται σε αυτό ποσοτικοποιείται με μια μέθοδο που βασίζεται στην ανάλυση του συγχρονισμού φάσης (Phase Dynamics Modelling, PDM). Η ιδέα του συγχρονισμού φάσεων προσφέρει μία νέα προοπτική στην κατανόηση και ποσοτικοποίηση των δυναμικών αλληλεπιδράσεων μεταξύ συστημάτων συζευγμένων ταλαντωτών. Λαμβάνοντας υπόψη την μη γραμμική φύση των νευρικών βιοσημάτων, παρουσιάζεται μια νέα προσέγγιση στην αναγνώριση της εγκεφαλικής δραστηριότητας και συσχέτισης κατά την διάρκεια πραγματικής και φανταστικής κίνησης του δείκτη των δακτύλων του χεριού. Χρησιμοποιήθηκε η τεχνική Empirical Mode Decomposition. Η τεχνική αυτή, η οποία είναι το βασικό αντικείμενο μελέτης της εργασίας ονομάζεται εμπειρικός τρόπος αποδόμησης του σήματος (empirical-mode-decomposition EMD) και χρησιμοποιείται για να γίνει αποσύνθεση των δεδομένων του ηλεκτροεγκεφαλογραφήματος (EEG) σε έναν αριθμό στοιχειωδών 3 ορθογώνιων συναρτήσεων (elementary – orthogonal modes) με καλά καθορισμένη χρονομεταβλητή συμπεριφορά.(IMFs). Πιθανό πλεονέκτημα αυτής της μεθόδου είναι ότι με τις παραδοσιακές, η μη γραμμικότητα του συστήματος ίσως οδηγήσει σε απώλεια της ενέργειας σε υψηλή συχνότητα με αποτέλεσμα να έχουμε ψεύτικο συγχρονισμό. Είναι μια νέα τεχνική μέθοδος ανάλυσης σημάτων η οποία βασίζεται στον μετασχηματισμό Hilbert – Huang. Σε αυτήν την εργασία δίνεται έμφαση στις ομοιότητες δομικής και λειτουργικής οργάνωσης των εγκεφαλικών διασυνδέσεων που πραγματοποιούνται κατά την διάρκεια πραγματικής και φανταστικής κίνησης. Στο πείραμα που μελετάται η ανάλυση συσχέτισης φάσης πραγματοποιείται στο κομμάτι (imf) που αντιστοιχεί στον βήτα ρυθμό (4th component). Η αναπαραγωγή των αποτελεσμάτων με την μέθοδο EMD στον βήτα ρυθμό και στην συνέχεια η ανάλυση συγχρονισμού φάσης προάγει αυτή την μέθοδο ως: προσαρμοστική (adaptive)-εναλλακτική (alternative) μέθοδο αναγνώρισης πραγματικών συναρτήσεων που προέρχονται από ολόκληρο το σήμα , παρέχοντας έναν νέο τρόπο καθορισμού συγκεκριμένου εύρους ζώνης συχνοτήτων μέσα στην οποία έχουμε φαινόμενα συσχέτισης. Η συνέπεια των αποτελεσμάτων τα οποία προέρχονται από ποικίλες καταγραφές αποτελούν σαφή ένδειξη ότι η ανάλυση του συγχρονισμού φάσης μπορεί να χρησιμοποιηθεί σαν ένα επιπλέον κριτήριο για ένα σύστημα (Ε-ΗΥ) προκειμένου να κωδικοποιήσει την πρόθεση της κίνησης.

Εφαρμογή της EMD στην ανάλυση δεδομένων που αφορούν την ηλιακή , παγκόσμια θερμοκρασία, τον αριθμό των ηλιακών κηλίδων και την συγκέντρωση του CO_2

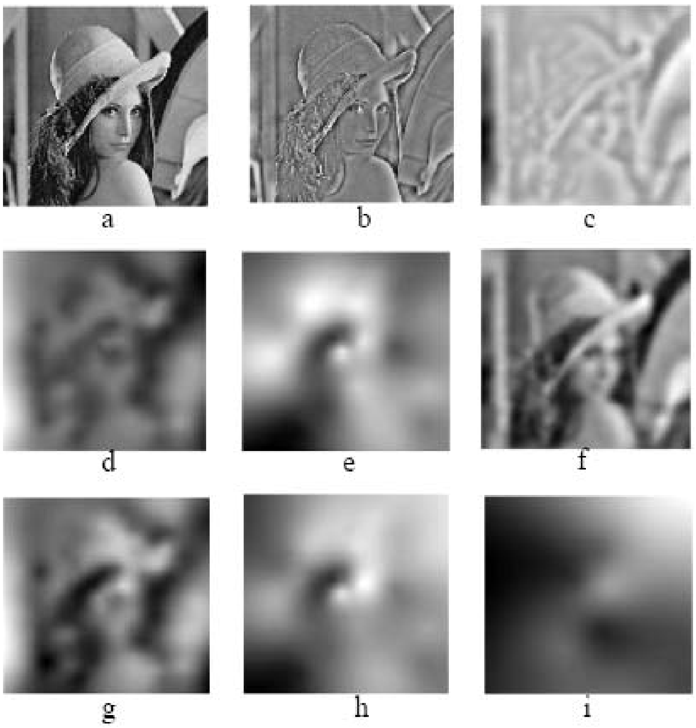
Οι διακυμάνσεις της συνολικής ηλιακής ακτινοβολίας, η παγκόσμια θερμοκρασία, ο αριθμός ηλιακών κηλίδων και η συγκέντρωση CO_2, αποσυντίθενται σε περιοδικά συστατικά χρησιμοποιώντας τη μέθοδο EMD. Τα περιοδικά συστατικά των δεδομένων αναλύονται και συγκρίνονται στον χρονικό ορίζοντα. Η 11χρονη περιοδική μεταβολή της παγκόσμιας μέσης θερμοκρασίας προσδιορίζεται και συγκρίνεται με την Schwabe cycle (ηλιακό κύκλο μαγνητικής δραστηριότητας) από τα προσεγγιστικά δεδομένα της ηλιακής ακτινοβολίας και των ηλιακών κηλίδων. Επίσης, η σχετική ακτινοβολούσα ισχύς των διαφορετικών περιοδικών χαρακτηριστικών της συνολικής ηλιακής ακτινοβολίας και της συγκέντρωσης του CO_2, εκτιμώνται εμπειρικά.

Ανάλυση των διακυμάνσεων της ροής ποταμού με τη χρήση EMD και την ανάλυση φάσματος με χρήση μετασχηματισμού Hilbert με αυθαίρετη σειρά

Σε αυτή την εργασία παρουσιάζουμε την ανάλυση δύο χρονοσειρών, που περιγράφουν την καθημερινή ροή δύο ποταμών. Η πρώτη περιλαμβάνει δεδομένα 32 χρόνων που αφορούν τον ποταμό Σηκουάνα (Γαλλία) και ο η δεύτερη τα δεδομένα 25 ετών που αφορούν τον ποταμό Wimereux (Γαλλία). Εφαρμόζουμε μια κλίμακα στις χρονοσειρές, με βάση την μέθοδο αποσύνθεσης EMD. Τα δεδομένα αποσυντίθενται σε πεπερασμένο αριθμό εγγενών λειτουργικών συναρτήσεων(IMFs). Η μέση συχνότητα κάθε IMF δείχνει ότι η EMD λειτουργεί ως μέσο αποθήκευσης των φιλτραρισμένων δεδομένων. Επιπλέον, η συσχέτιση των IMFs των ποταμών Σηκουάνα και Wimereu υποδεικνύει συσχέτιση μεταξύ των τμημάτων ευρείας κλίμακας των IMFs, γεγονός που δείχνει ότι και οι δύο ποταμοί είναι πιθανόν να επηρεαστούν από την ίδια θαλάσσια κλιματική εκδήλωση στην Βόρεια Γαλλία. Ως επιβεβαίωση, διαπιστώσαμε ότι τα τμήματα ευρείας κλίμακας έχουν την ίδια τάση εξέλιξης. Θα εφαρμοστεί τελικά ανάλυση φάσματος με χρήση μετασχηματισμού Hilbert, μια τεχνική που προέκυψε από την ανάλυση χρονοσειρών και τις μελέτες αναταραχής που αφορούν την εκροή του ποταμού Σηκουάνα. Αυτή η μέθοδος παρέχει μια αναπαράσταση ευρέων συχνοτήτων, της αρχικής χρονοσειράς. Όταν υπολογίζονται οι οριακές τιμές του εύρους, μπορούμε να παρατηρήσουμε μια διαλείπουσα μελέτη στη συχνότητα του χώρου. Εφαρμοσμένη στα δεδομένα εκροής του ποταμού Σηκουάνα, αυτό δείχνει την κλιμάκωση του εύρους και χαρακτηρίζει τις διακοπτόμενες διακυμάνσεις στο εύρος των κλιμάκων από 4,5 έως 60 μέρες, μεταξύ συνοπτικής και εποχιακής κλίμακας.

Εφαρμογή της EMD στην ανάλυση εικόνων

Στόχος της συγκεκριμένης εργασίας ήταν να μειωθούν οι παράμετροι που αντιπροσωπεύουν την εικόνα χρησιμοποιώντας την μέθοδο EMD.
Στην παρακάτω εικόνα παρουσιάζεται η ανάλυση της φωτογραφίας όπως προέκυψε με την μέθοδο EMD:



Εικόνα 3

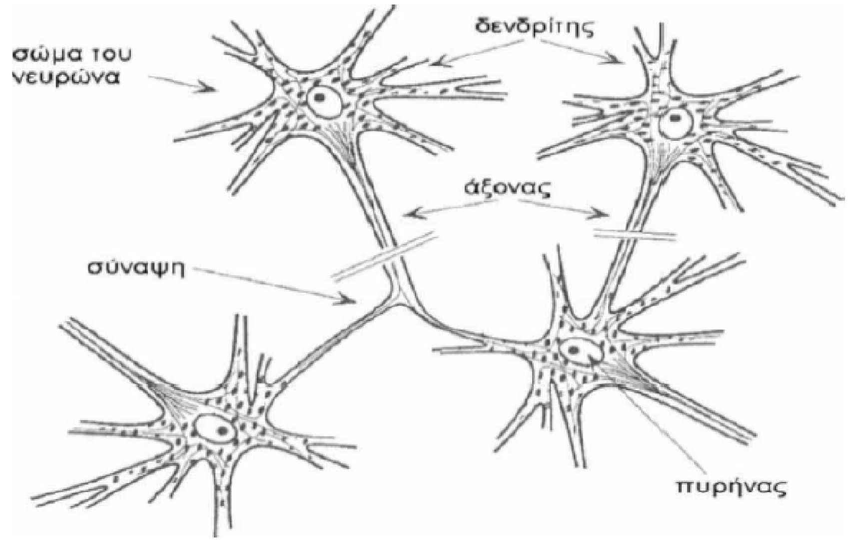
Όπου:
a)Αρχική εικόνα
b) Πρώτο IMF
c) Δεύτερο IMF
d) Τρίτο IMF
e) Τέταρτο IMF
f) Πρώτο υπόλειμμα
g) Δεύτερο υπόλειμμα
h) Τρίτο υπόλειμμα
i) Τέταρτο υπόλειμμα.

ΚΕΦΑΛΑΙΟ 3: Νευρωνικά Δίκτυα

3.1. Εισαγωγή

Το ανθρώπινο μυαλό έχει την ικανότητα να αποθηκεύει μεγάλη ποσότητα αλλά και ποικιλία γνώσης, να συσχετίζει πληροφορίες με μεγάλη ταχύτητα και ευκολία καθώς και να προσαρμόζεται στις συνθήκες και στα προβλήματα που αντιμετωπίζει. Ο ανθρώπινος εγκέφαλος δηλαδή, έχει την ικανότητα μάθησης λόγω εμπειρίας. Αυτό οφείλεται στην αρχιτεκτονική λειτουργίας που χρησιμοποιεί. Πιο συγκεκριμένα, ο εγκέφαλος περιλαμβάνει χιλιάδες νευρώνες, διατεταγμένους με τέτοιο τρόπο ώστε να μπορούν να επεξεργάζονται πολλές πληροφορίες ταυτόχρονα. Σε αυτήν ακριβώς την ικανότητα του εγκεφάλου βασίστηκαν οι έρευνες γύρω από τα Τεχνητά Νευρωνικά Δίκτυα.

Η δομή ενός βιολογικού νευρώνα παρουσιάζεται στην παρακάτω εικόνα:



<u>Σχήμα 2: Απλοποιημένη απεικόνιση τεσσάρων νευρώνων</u>

Από την εικόνα μπορούμε να διακρίνουμε τα τρία κύρια μέρη από τα οποία αποτελείται ένας νευρώνας:

- Το κυτταρικό σώμα
- τους δενδρίτες
- και τους άξονες

Το κυτταρικό σώμα αποτελεί τον πυρήνα του νευρώνα. Οι δενδρίτες είναι πολλές δενδροειδής αποφύσεις, οι οποίες λαμβάνουν ηλεκτρικά σήματα, που μεταφέρονται μέσω του άξονα σε άλλους νευρώνες. Το σημείο σύνδεσης του άξονα με τους δενδρίτες ονομάζεται *σύναψη*. Τα σήματα εισόδου από τους νευρώνες μεταφέρονται μέσω ενισχυτικών, είτε κατασταλτικών συνάψεων, οι οποίες είτε αυξάνουν είτε μειώνουν αντίστοιχα την πιθανότητα διέγερσης ενός νευρώνα. Η ικανότητα μάθησης και μνήμης που παρουσιάζει ο εγκέφαλος οφείλεται στην ικανότητα των συνάψεων να μεταβάλουν την αγωγιμότητά τους. Έπειτα, τα σήματα προωθούνται από τους δενδρίτες στο σώμα όπου αθροίζονται, και αν το άθροισμα είναι μεγαλύτερο από το δυναμικό ενεργοποίησης του άξονα, τότε ο νευρώνας διεγείρεται, Στη συνέχεια, τα άκρα του άξονα στέλνουν το σήμα σε άλλη σύναψη και ο νευρώνας απενεργοποιείται έτσι ώστε να είναι σε θέση ενεργοποίησης από άλλο σήμα.

Τα νευρωνικά δίκτυα αποτελούνται από πολλούς διασυνδεδεμένους νευρώνες με παρόμοια χαρακτηριστικά, όπως εισόδους, εξόδους, βάρη ενεργοποίησης και τάση πόλωσης. Κάθε νευρωνικό δίκτυο έχει διαφορετικές ιδιότητες. Δηλαδή μαθαίνει, θυμάται, συσχετίζει, συγκρίνει τη νέα πληροφορία με την υπάρχουσα γνώση, ταξινομεί τις νέες πληροφορίες, με διαφορετικό τρόπο και με διαφορετική ταχύτητα από τα άλλα.

Τα νευρωνικά δίκτυα δεν επεξεργάζονται την πληροφορία σειριακά, αλλά με την μέθοδο της Παράλληλης Ανάλυσης της σύνθετης πληροφορίας στα βασικά της στοιχεία.

Για παράδειγμα ένα βασικό χρώμα, μπορεί να διαχωριστεί στις βασικές του συχνότητες και πλάτη, και ανά πάσα στιγμή να δημιουργηθεί ξανά από αυτές. Ομοίως ένα νευρωνικό δίκτυο αποσυνθέτει την σύνθετη πληροφορία στα βασικά της στοιχεία. Αυτά τα στοιχεία και η μεταξύ τους συσχέτιση αποθηκεύονται σε ένα μέρος του εγκεφάλου. Για παράδειγμα όταν ένας άνθρωπος βλέπει μια εικόνα δεν αποθηκεύει στη μνήμη του έναν πίνακα από εικονοστοιχεία όπως γίνεται στην περίπτωση του ηλεκτρονικού υπολογιστή, αλλά αποθηκεύει στοιχειώδη χαρακτηριστικά όπως γραμμές, σχήματα και χρώματα.

3.2. Τεχνητά Νευρωνικά Δίκτυα

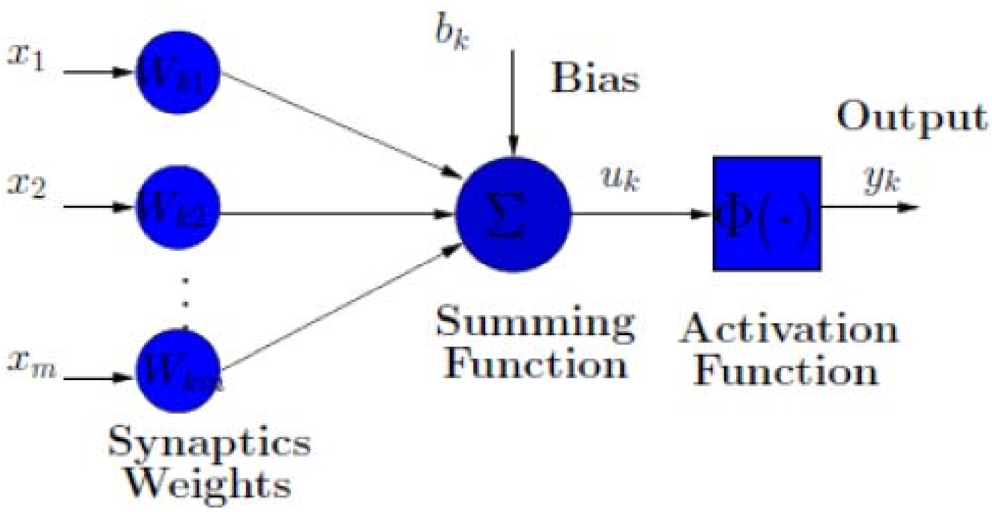
Το τεχνητό νευρωνικό δίκτυο είναι ένα υπολογιστικό μοντέλο που βασίζεται στο μοντέλο του κεντρικού ανθρώπινου νευρικού συστήματος. Οι νευρώνες είναι τα δομικά στοιχεία του δικτύου και αλληλοσυνδέονται με τέτοιο τρόπο ώστε να έχουν την ικανότητα επεξεργασίας πληροφοριών ταυτόχρονα.

Τα τεχνητά νευρωνικά δίκτυα χωρίζονται σε δύο κατηγορίες:

- Τα στατικά νευρωνικά δίκτυα, τα οποία δεν περιέχουν στοιχεία με μνήμη αλλά μπορούν να δέχονται σαν εισόδους προηγούμενες τιμές των εισόδων.
- Δυναμικά νευρωνικά δίκτυα, τα οποία διαθέτουν στοιχεία με μνήμη και είναι κατάλληλα για την προτυποποίηση μη γραμμικών δυναμικών συστημάτων.

3.3. Βασικό Μοντέλο Νευρώνα

Στην παρακάτω εικόνα παρουσιάζεται ένα μοντέλο τεχνητού νευρωνικού δικτύου.



Σχήμα 3: Νευρωνικό δίκτυο

Ένα τεχνητό νευρωνικό δίκτυο δέχεται κάποια σήματα εισόδων x_0, x_1, x_2, ……, x_m , καθένα από τα οποία μεταβάλλεται από μία τιμή βάρους w_i , ο ρόλος του οποίου είναι αντίστοιχος της σύναψης του βιολογικού εγκεφάλου. Η τιμή του βάρους αναπαριστά στην διέγερση ή την αποδιέγερση της εισόδου του νευρώνα και είναι θετική ή αρνητική αντίστοιχα. Όπως φαίνεται και στην εικόνα το μοντέλο του νευρώνα παρουσιάζει και ένα εξωτερικό βάρος b_k που έχει σαν αποτέλεσμα την αύξηση ή την μείωση της τιμής που δίνεται σαν είσοδος στο δίκτυο στην συνάρτηση ενεργοποίησης ανάλογα με το αν είναι θετικό ή αρνητικό. Η τιμή της εισόδου του είναι πάντα η μονάδα, $x_0 = 1$. Η απόλυτη τιμή του βάρους καθορίζει την ισχύ της σύνδεσης.

Το σώμα του νευρώνα χωρίζεται σε δύο μέρη, τον αθροιστή και την συνάρτηση ενεργοποίησης. Ο αθροιστής προσθέτει τα πολλαπλασιασμένα με τα αντίστοιχα βάρη, σήματα εισόδων. Η συνάρτηση ενεργοποίησης είναι ένα μη γραμμικό φίλτρο το οποίο διαμορφώνει την τελική τιμή y του σήματος, σε συνάρτηση με την ποσότητα u_k

Επομένως οι συναρτήσεις που περιγράφουν το μοντέλο ενός στατικού νευρώνα είναι οι εξής:

$$u_k = \sum_{l=1}^{m} w_{kl}\, x_l$$

και η έξοδος y_k δίνεται:

$$\Phi\left\{\sum_{l=1}^{m} w_{kl}x_l + b_k\right\} \text{ ή } \Phi\left\{\sum_{l=1}^{m+1} w_{kl}x_l\right\} \quad \text{με } x_{m+1} = 1 \text{ και } b_k = w_{k\,m+1}$$

όπου έχουμε:

εισερχόμενα σήματα: $x_1, x_2, \ldots, x_m$
τα βάρη των συνάψεων του νευρώνα k: $W_{k1}, W_{k2}, \ldots, W_{km}$
η έξοδος του γραμμικού συνδυαστή: u_k
η συνάρτηση ενεργοποίησης: $\Phi(.)$
το σήμα που δίνει σαν έξοδο ο νευρώνας k: y_k

Οι συντελεστές διασύνδεσης ή συναπτικά βάρη του δικτύου, είναι αρχικά άγνωστα και στόχος της μάθησης είναι η συστηματική εύρεσή τους ελαχιστοποιώντας κάποιο μέτρο του σφάλματος μεταξύ της επιθυμητής και της πραγματικής εξόδου του δικτύου.

Αν ο νευρώνας είναι δυναμικός, δηλαδή αν έχει μνήμη, τότε η είσοδος της συνάρτησης ενεργοποίησης δεν εξαρτάται μόνο από το σταθμισμένο άθροισμα αλλά και από την προηγούμενη τιμή του.

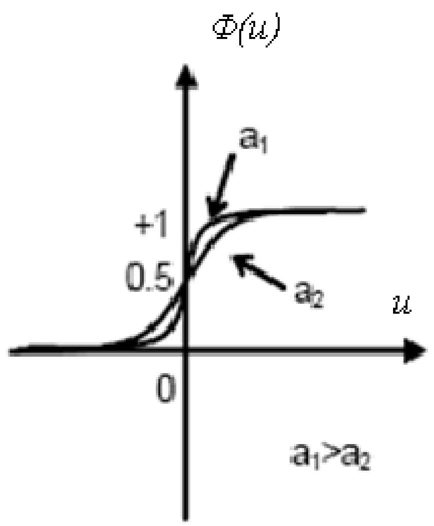
$$u_k = u_{k-1} + \sum_{l=1}^{m} w_{kl}\, x_l$$

Όπου k ο δείκτης του χρόνου δειγματοληψίας. Οι δυναμικοί νευρώνες είναι ιδιαίτερα χρήσιμοι για την ταυτοποίηση δυναμικών διαδικασιών και την προσομοίωση δυναμικών συστημάτων.

Όταν η τιμή του σταθμισμένου βάρους υπερβεί την τιμή πόλωσης (bias) b, ο νευρώνας ενεργοποιείται. Το σταθμισμένο άθροισμα που υπολογίζεται, χρησιμοποιείται σαν είσοδος της συνάρτησης ενεργοποίησης. Όπως αναφέρθηκε παραπάνω, η συνάρτηση ενεργοποίησης πρέπει να είναι ένα μη γραμμικό φίλτρο ώστε να μοντελοποιεί μη γραμμικά φαινόμενα. Έχουν προταθεί πολλές συναρτήσεις ενεργοποίησης. Η πιο ευρέως χρησιμοποιούμενη είναι η σιγμοειδής ή λογιστική συνάρτηση.

Η λογιστική (sigmoid), μέλος της οικογένειας των σιγμοειδών:

$$\Phi(u) = \frac{1}{1+e^{-au}}$$

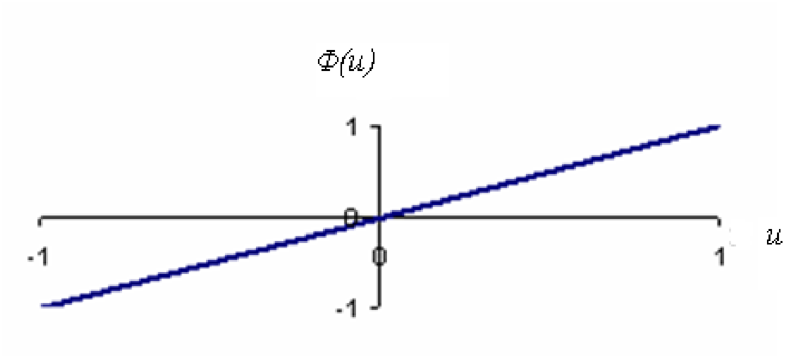


<u>Σχήμα 4: Λογιστική συνάρτηση</u>

Παρακάτω ακολουθούν μερικές ακόμα συναρτήσεις ενεργοποίησης:

- Η γραμμική (Linear):

  $\Phi(u)=u$



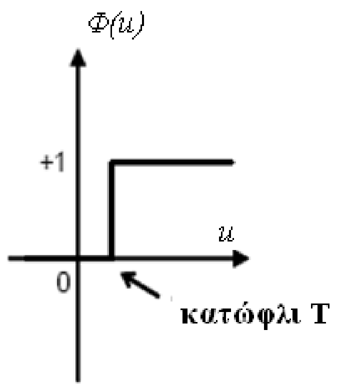
Σχήμα 5: Γραμμική συνάρτηση

- Η σχέση με υπερβολική εφαπτομένη :

$$\Phi(u) = tansig(u) = \frac{1 - e^{-u}}{1 + e^{-u}}$$

- Η **βηματική (step) ή κατωφλίου συνάρτηση:**

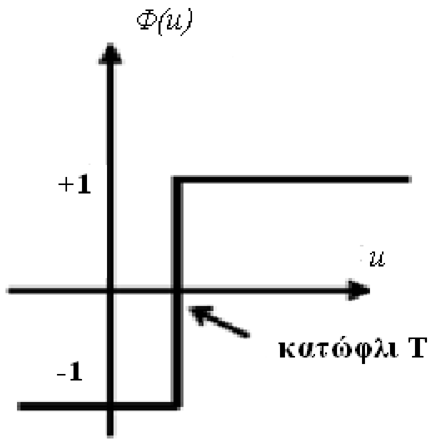
$$\Phi(u) = \begin{cases} 1, & u \geq 0 \\ 0, & u < 0 \end{cases}$$



Σχήμα 6: Βηματική συνάρτηση

- Η συνάρτηση ***πρόσημου (sign)***

$$\Phi(u) = \begin{cases} 1, & u \geq T \\ -1, & u < T \end{cases}$$



Σχήμα 7: Συνάρτηση προσήμου

Η συνάρτηση ημίτονου ή συνημίτονου:

$$\Psi(u) = \sin(u) \text{ ή } \Psi(u) = \cos(u)$$

Η σχέση τύπου Percepton:

$$f(\sigma) = \begin{cases} \sigma, & \varepsilon\alpha\nu\ \sigma > 0 \\ 0, & \varepsilon\alpha\nu\ \sigma < 0 \end{cases}$$

3.4. Τοπολογία νευρωνικών δικτύων

Τα τεχνητά νευρωνικά δίκτυα αποτελούνται από συμπλέγματα νευρώνων δομημένους κατά στρώματα ή επίπεδα (layers). Μπορεί να αποτελούνται είτε από ένα στρώμα (μονοστρωματικά δίκτυα) είτε από περισσότερα (πολυστρωματικά δίκτυα).

Τα τεχνητά νευρωνικά δίκτυα χωρίζονται σε δύο κατηγορίες, σύμφωνα με την τοπολογία τους:

3.4.1. Διαδιδόμενα προς τα εμπρός ή πρόσω τροφοδοτούμενα ΤΝΔ (FeedForward):

Οι μονάδες στέλνουν τις εξόδους τους σε μονάδες, από τις οποίες δεν λαμβάνουν είσοδο αμέσως ή εμμέσως. Αυτό έχει σαν αποτέλεσμα να μην υπάρχουν βρόγχοι ανάδρασης και τα δίκτυα να είναι οργανωμένα σε επίπεδα, των οποίων οι μονάδες είναι συνδεδεμένες μόνο με τις μονάδες των επόμενων επιπέδων (αυστηρά πρόσω τροφοδοτούμενο ΤΝΔ). Αριθμώντας κατάλληλα τις μονάδες μπορούμε να ορίσουμε έναν πίνακα βαρών **W** ο οποίος είναι κάτω-τριγωνικός και έχει μηδενική διαγώνιο (εφόσον η μονάδα δεν ανατροφοδοτεί τον εαυτό της).

3.4.1.1. Μονοστρωματικό Δίκτυο Τύπου Percepton

Το πιο απλό είδος νευρωνικού δικτύου είναι το μονοστρωματικό δίκτυο τύπου Percepton. Σε αυτό τον τύπο δικτύων υπολογίζεται το σταθμισμένο άθροισμα των εισόδων σε κάθε κόμβο. Αν η τιμή του αθροίσματος ξεπεράσει την τιμή ενός κατωφλίου , η οποία έχει οριστεί εκ των προτέρων, τότε ο νευρώνας ενεργοποιείται. Στη βιβλιογραφία κατά κύριο λόγο ο παραπάνω όρος αναφέρεται σε δίκτυα με έναν κόμβο.

3.4.1.2. Πολυστρωματικά Δίκτυα (Multy-layer Networks).

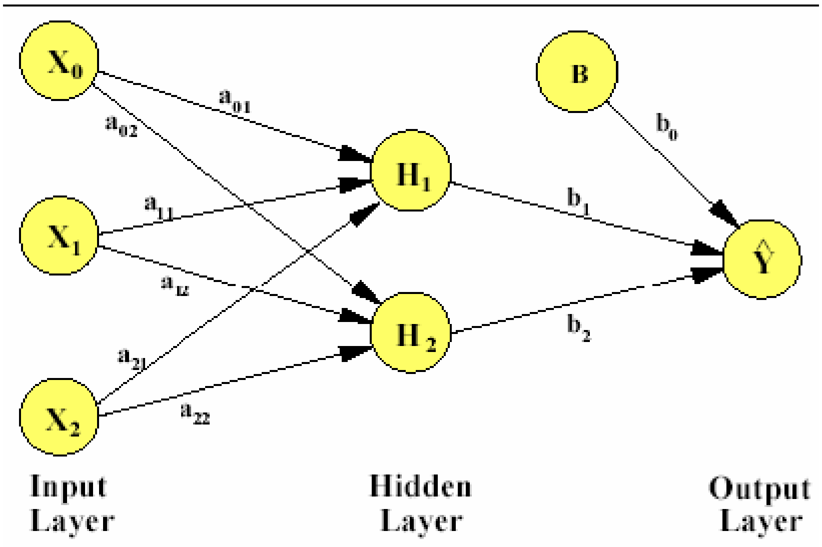
Στο συγκεκριμένο τύπο δικτύων έχουμε πολλά επίπεδα και η ροη της πληροφορίας γίνεται πάλι μόνο προς τα εμπρός , δηλαδή από το στρώμα εισόδου προς το στρώμα εξόδου δια μέσω των κρυφών- ενδιάμεσων στρωμάτων, αν υπάρχουν τέτοια, χωρίς καμία ανατροφοδότηση.

3.4.1.3. Νευρωνικά Δίκτυα Με Κρυμμένα Επίπεδα

Τα δίκτυα που περιγράφηκαν μέχρι στιγμής είχαν μια πολύ απλή δομή δύο επιπέδων που συνέδεαν τις εισόδους με τις εξόδους. Στις πραγματικές εφαρμογές, η αρχιτεκτονική του δικτύου είναι πιο πολύπλοκη. Οι ερευνητές σχεδόν πάντα σχεδιάζουν δομές που περιλαμβάνουν ένα ή περισσότερα *κρυμμένα επίπεδα*, όπως στο σχήμα 8. Σ' αυτό το σχήμα τα a_{ij} δηλώνουν τα βάρη που συνδέουν την είσοδο *i* με την κρυμμένη μονάδα *j*. Υποθέτουμε ότι X_0 είναι ο όρος σταθερού βάρους (δηλαδή ο συντελεστής ανάσχεσης) για τις κρυμμένες μονάδες, δηλαδή X_0=1 , ενώ το B είναι ο όρος σταθερού βάρους για τη μονάδα εξόδου. Σε αντίθεση με τις μονάδες εισόδου και εξόδου, οι κρυμμένες μονάδες δεν αναπαριστούν κάποια πραγματική έννοια. Δεν έχουν ούτε ερμηνεία ούτε νόημα. Είναι απλά το ενδιάμεσο αποτέλεσμα στη διαδικασία υπολογισμού της τιμής εξόδου. Συνεπώς δεν έχουν αντίστοιχη έννοια στην οικονομετρία. Οι κρυμμένες μονάδες συμπεριφέρονται όπως οι μονάδες εξόδου, δηλαδή υπολογίζουν το σταθμικό άθροισμα των μεταβλητών εισόδου και έπειτα επεξεργάζονται το αποτέλεσμα χρησιμοποιώντας τη συνάρτηση ενεργοποίησης, που είναι σχεδόν πάντα η συνάρτηση logistic. Στο δίκτυο που απεικονίζεται στο σχήμα 8, το αποτέλεσμα που παράγουν οι κρυμμένες μονάδες θα είναι:

$$H_1 = f(a_{01} + a_{11}X_1 + a_{21}X_2) = \frac{1}{1 + e^{-(a_{01} + a_{11}X_1 + a_{21}X_2)}}$$

$$H_2 = f(a_{02} + a_{21}X_1 + a_{22}X_2) = \frac{1}{1 + e^{-(a_{02} + a_{21}X_1 + a_{11}X_1)}}$$



Σχήμα 8: Το feedforward νευρωνικό δίκτυο με ένα κρυμμένο επίπεδο

Τοποθετώντας τη συνάρτηση ενεργοποίησης logistic στις κρυμμένες μονάδες και όχι στη μονάδα εξόδου, το δίκτυο δεν περιορίζεται πλέον στην παραγωγή εκτιμήσεων πεπερασμένων μεταβλητών. Εάν η εξαρτημένη μεταβλητή είναι μη πεπερασμένη, η μονάδα εξόδου συνήθως θα χρησιμοποιήσει ως συνάρτηση ενεργοποίησης την ταυτότητα, δηλαδή η έξοδος θα ισούται με το σταθμικό άθροισμα των τιμών των κρυμμένων μονάδων, σταθμισμένο από τους συντελεστές b_j. Αυτό συνεπάγεται συνεχή, μη γραμμική, μη πεπερασμένη έξοδο όπως εκφράζεται στην εξίσωση που ακολουθεί:

$$Y = b_0 + b_1 H_1 + b_2 H_2$$

$$Y = b_0 + \frac{b_1}{1 + e^{-(a_{01} + a_{11}X_1 + a_{21}X_2)}} + \frac{b_2}{1 + e^{-(a_{02} + a_{12}X_1 + a_{22}X_2)}}$$

Εάν η εξαρτημένη μεταβλητή είναι πεπερασμένη, η έξοδος χρησιμοποιεί συνήθως τη συνάρτηση ενεργοποίησης logistic, παράγοντας έτσι μια πεπερασμένη έξοδο, όπως στη παρακάτω σχέση :

$$Y = f(b_0 + b_1H_1 + b_2H_2)$$

$$Y = \frac{1}{1 + e^{-(b_0 + b_1H_1 + b_2H_2)}}$$

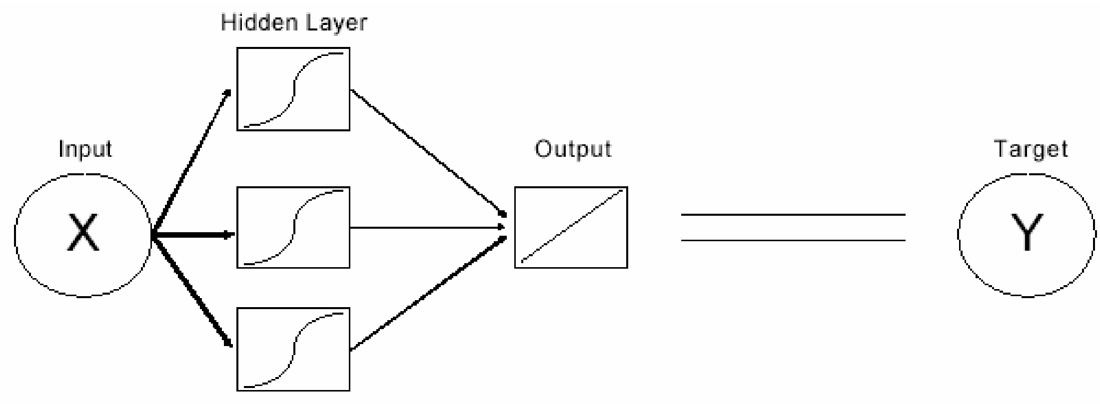
$$Y = \frac{1}{1 + e^{-(b_0 + \frac{b_1}{1+e^{-(a_{01} + a_{11}X! + a_{21}X_2)}} + \frac{b_2}{1+e^{-(a_{02} + a_{12}X_1 + a_{22}X_2)}}}}$$

Συνεπώς, το δίκτυο έχει τη δυνατότητα παραγωγής πεπερασμένης ή μη εξόδου, διατηρώντας ταυτόχρονα τα μη γραμμικά χαρακτηριστικά του. Ο συνυπολογισμός κρυμμένων μονάδων στο δίκτυο προσφέρει ένα ακόμη σημαντικό πλεονέκτημα. Πολλοί ερευνητές απέδειξαν πέραν κάθε αμφιβολίας ότι ένα νευρωνικό δίκτυο τριών επιπέδων με τη συνάρτηση ενεργοποίησης logistic στις κρυμμένες μονάδες, είναι *μια γενική εκτιμήτρια (universal approximator)*. Αυτό σημαίνει ότι εάν συμπεριληφθεί ένας σημαντικός αριθμός κρυμμένων μονάδων, το δίκτυο είναι σε θέση να προσδιορίσει σχεδόν οποιαδήποτε γραμμική ή μη συνάρτηση σε ένα επιθυμητό επίπεδο ακριβείας. Επομένως τα νευρωνικά δίκτυα δύναται να χρησιμοποιηθούν ως ισχυρό εργαλείο για την αναγνώριση και την αναπαραγωγή πολύπλοκων μη γραμμικών δεδομένων σε χρονοσειρές. Είτε μελετάμε το πραγματικό προϊόν (GDP), είτε τον πληθωρισμό, είτε την ανεργία, είτε την τιμή συναλλάγματος, θεωρητικά ένα νευρωνικό δίκτυο πρέπει να είναι ικανό να ανιχνεύσει και να αντιγράψει οποιοδήποτε πολύπλοκο μη γραμμικό υπόδειγμα από τα δεδομένα. Επιπλέον, δεν απαιτείται *a priori* γνώση της διαδικασίας δημιουργίας των δεδομένων, όπως συμβαίνει στα μοντέλα μη γραμμικής παλινδρόμησης. Αρκεί μόνο η χρήση μιας γενικής συναρτησιακής μορφής, αλλά με μεγαλύτερο αριθμό κρυμμένων μονάδων.

Όπως αναφέρθηκε, δεν υπάρχει θεωρητική βάση για τον προσδιορισμό του αριθμού των κρυμμένων μονάδων ή επιπέδων στο δίκτυο. Βάσει της ιδιότητας γενικής εκτιμήτριας που περιγράφηκε παραπάνω, είναι λογική η χρήση μεγάλου αριθμού κρυμμένων μονάδων. Εντούτοις, εάν προστεθούν πολλά κρυμμένα επίπεδα, το δίκτυο μπορεί να γίνει επιρρεπές στο πρόβλημα της υπερεκπαίδευσης (*overfitting*) . Αυτό σημαίνει ότι το δίκτυο μπορεί να πετύχει μια υψηλότερη ακρίβεια πρόβλεψης για την περίοδο που εκτιμάται, αλλά θα παράγει φτωχές προβλέψεις όταν γίνει χρήση δεδομένων εκτός του δείγματος εκπαίδευσης. Επιπλέον, το πλήθος των βαρών του δικτύου αυξάνεται ραγδαία, καθώς προστίθενται περισσότερες κρυμμένες μονάδες, αυξάνοντας έτσι το χρόνο εκτίμησης του μοντέλου. Στην πράξη, η σχεδίαση της αρχιτεκτονικής του δικτύου είναι μια μονότονη διαδικασία δοκιμής και σφάλματος. Οι ερευνητές συνήθως εκτιμούν πολλά διαφορετικά δίκτυα και επιλέγουν εκείνο που οδηγεί στο μικρότερο σφάλμα πρόβλεψης.

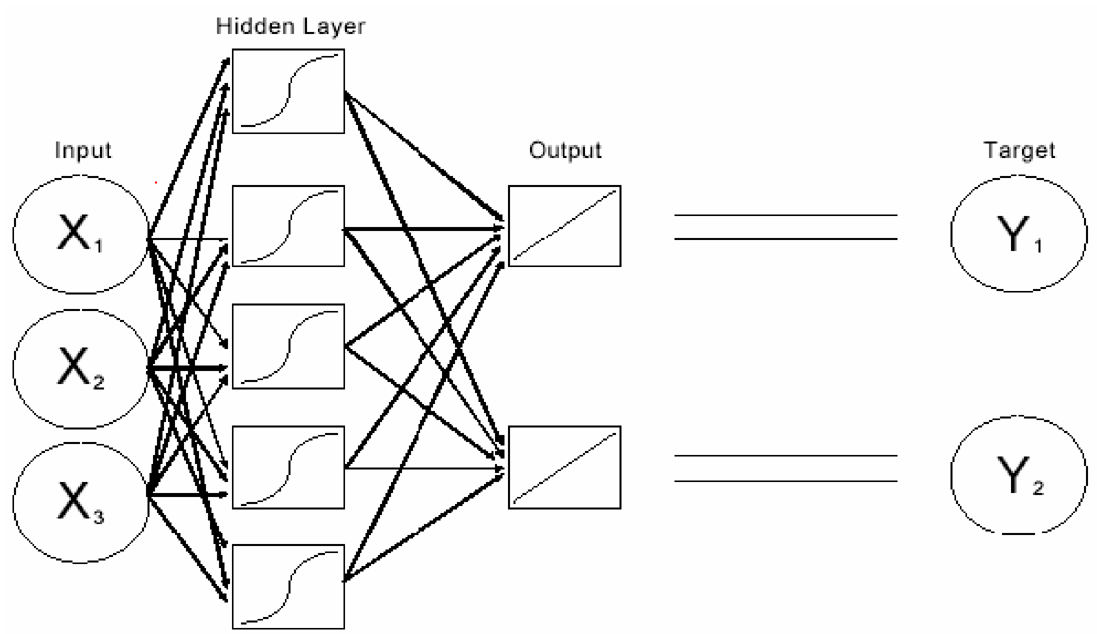
3.1.1.4. Νευρωνικά Δίκτυα Πολλαπλών Επιπέδων (MLP)

Μία από τις χρησιμότερες και πλέον επιτυχημένες εφαρμογές των νευρωνικών δικτύων στην ανάλυση δεδομένων και στην πρόβλεψη αποτελεί το μοντέλο πολλαπλών επιπέδων percepton (multilayer perceptron model) MLP. Τα μοντέλα πολλαπλών επιπέδων perceptron είναι μη γραμμικά μοντέλα νευρωνικών δικτύων που μπορούν να χρησιμοποιηθούν σχεδόν με όλες τις συναρτήσεις με υψηλό βαθμό ακρίβειας. Τα MLP περιέχουν ένα ή περισσότερα κρυμμένα επίπεδα νευρώνων που χρησιμοποιούν μη γραμμικές συναρτήσεις ενεργοποίησης, όπως συναρτήσεις logistic. Στο σχήμα 2.8 απεικονίζεται ένα MLP με ένα κρυμμένο επίπεδο και μία μόνο είσοδο και έξοδο. Το MLP στο σχήμα 2.8 αναπαριστά μια απλή μη γραμμική παλινδρόμηση.



Σχήμα 9: Perceptron Πολλαπλών επιπέδων με μία μόνο είσοδο και έξοδο

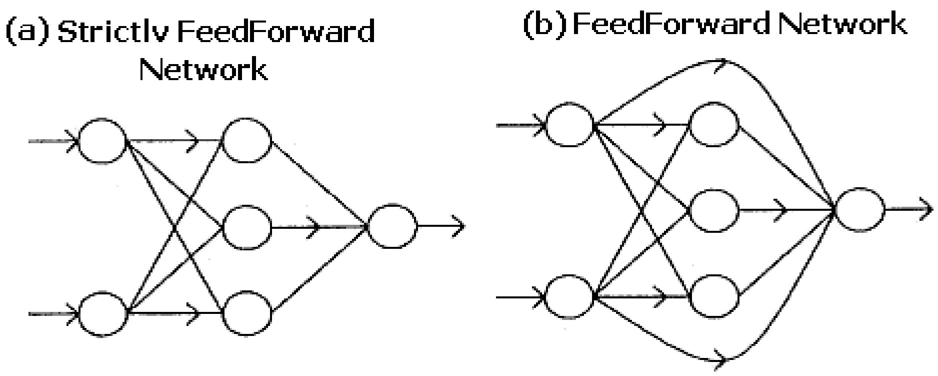
Ανάλογα με τους τύπους των δεδομένων, ο αριθμός εισόδων-εξόδων στο MLP μεταβάλλεται. Στο παρακάτω σχήμα απεικονίζεται μια πολλαπλή μη γραμμική παλινδρόμηση πολλαπλών μεταβλητών. Η πολυπλοκότητα των MLP μπορεί να προσαρμοσθεί ανάλογα με τον αριθμό των κρυμμένων επιπέδων. Ανάλογα με τον αριθμό των κρυμμένων επιπέδων, ένα MLP μετατρέπεται από απλό παραμετρικό μοντέλο σε ένα ευέλικτο μη παραμετρικό μοντέλο.



Σχήμα 10: Perceptron πολλαπλών επιπέδων με πολλαπλές εισόδους και εξόδους

Ένα MLP αποτελείται συνήθως από πολλά επίπεδα κόμβων. Το πρώτο (χαμηλότερο), είναι το επίπεδο εισόδου όπου λαμβάνεται η εξωτερική πληροφορία. Το τελευταίο (υψηλότερο), είναι το επίπεδο εξόδου όπου επιτυγχάνεται η λύση του προβλήματος. Ενδιάμεσα υπάρχουν τα κρυμμένα επίπεδα. Οι κόμβοι στα γειτονικά επίπεδα συνήθως είναι πλήρως συνδεδεμένοι μέσω τόξων από χαμηλότερα επίπεδα προς υψηλότερα.

Παρακάτω φαίνονται παραδείγματα δικτύων **FF**, και αυστηρά **FF**.



<u>Σχήμα 11: a) εμπρός τροφοδοτούμενο ΤΝΔ,b) Αυστηρώς εμπρός τροφοδοτούμενο ΤΝΔ</u>

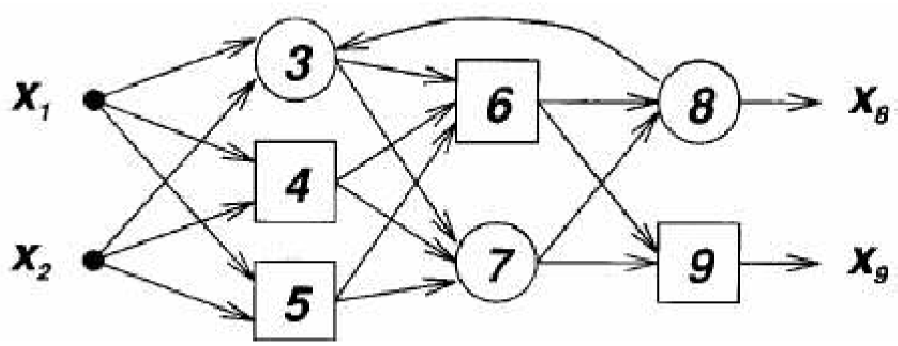
Ένα από τα πλεονεκτήματα ενός FF ΤΝΔ είναι ότι εφαρμόζει στατική αντιστοιχία από το χώρο των εισόδων του στο χώρο των εξόδων του σε αντίθεση με ένα FB ΤΝΔ που είναι ένα μη γραμμικό δυναμικό σύστημα.

3.4.2. Τεχνητά Νευρωνικά Δίκτυα ανατροφοδότησης (FeedBack):

Η διαφορά ενός τέτοιου δικτύου με ένα FeedForward ΤΝΔ είναι ότι επιτρέπεται η ύπαρξη βρόγχων ανάδρασης. Αυτό σημαίνει ότι σε έναν τουλάχιστον νευρώνα, το σήμα εξόδου του επηρεάζει το σήμα που έρχεται στην είσοδο του νευρώνα. Για παράδειγμα, ένα ΑΤΝΔ μπορεί να αποτελείται από ένα μόνο επίπεδο νευρώνων όπου κάθε νευρώνας επιστρέφει το σήμα εξόδου του σαν σήμα εισόδου σε όλους τους άλλους νευρώνες του επιπέδου.

Αντίστοιχα στα πολυστρωματικά δίκτυα με ανατροφοδότηση υπάρχει τουλάχιστον ένας κλειστός βρόχος (closed loop) από τον οποίο δεδομένα από ανώτερα στρώματα τροφοδοτούνται στα κατώτερα στρώματα. Τέτοιου είδους δίκτυα είναι τα δίκτυα Hopfield στα οποία όλες οι συνδέσεις είναι συμμετρικές.

Παρακάτω παρουσιάζεται η μορφή ενός νευρωνικού δικτύου με ανατροφοδότηση.



Σχήμα 12: Δίκτυο με ανατροφοδότηση

3.5 . Μεθοδολογία κατασκευής τεχνητών νευρωνικών δικτύων

Πριν σχεδιάσουμε ένα τεχνητό νευρωνικό δίκτυο, πρέπει να λάβουμε υπόψη πολλά ζητήματα που αφορούν την μοντελοποίηση του δικτύου, γιατί αυτά επηρεάζουν την απόδοσή του. Πιο συγκεκριμένα πρέπει να προσδιορίσουμε τα εξής:

- την αρχιτεκτονική του δικτύου
- την συνάρτηση ενεργοποίησης
- τον αλγόριθμο εκπαίδευσης

3.5.1. Αρχιτεκτονική δικτύου

Τα τεχνητά νευρωνικά δίκτυα αποτελούνται από επίπεδα και κόμβους. Όσον αφορά το δίκτυο πολλαπλών επιπέδων (MLP), αποτελείται από το επίπεδο εισόδου, που περιλαμβάνει τους κόμβους εισόδου, από το επίπεδο εξόδου, που περιλαμβάνει τους κόμβους εξόδου, και ενδιάμεσα υπάρχουν ένα ή περισσότερα κρυμμένα επίπεδα που περιλαμβάνουν τους κρυμμένους κόμβους. Τα βήματα που πρέπει να ακολουθούνται για την κατασκευή ενός δικτύου πολλαπλών επιπέδων είναι τα εξής:

- προσδιορισμός των κόμβων εισόδου
- προσδιορισμός των κρυμμένων επιπέδων και κόμβων, και
- προσδιορισμός των κόμβων εξόδου

Αν και υπάρχουν πολλές διαφορετικές προσεγγίσεις για την εύρεση της βέλτιστης αρχιτεκτονικής τεχνητού νευρωνικού δικτύου, καμία από αυτές δεν μπορεί να εγγυηθεί τη βέλτιστη λύση για όλα τα πραγματικά προβλήματα πρόβλεψης. Ενδεικτικά, κάποιες από αυτές τις προσεγγίσεις είναι ο αλγόριθμος πολυωνυμικού χρόνου, ο αλγόριθμος κλαδέματος, η τεχνική κανονικής αποσύνθεσης αποσύνθεσης (canonical decomposition technique) και το κριτήριο πληροφορίας του δικτύου.

3.5.1.1. Ο αριθμός των κρυμμένων επιπέδων και κόμβων

Οι κρυμμένοι κόμβοι στο κρυμμένο επίπεδο επιτρέπουν στα νευρωνικά δίκτυα να ανιχνεύουν το χαρακτηριστικό, να συλλαμβάνουν το υπόδειγμα στα δεδομένα και να εκτελούν μη γραμμικές συναρτήσεις μεταξύ των μεταβλητών εισόδου και εξόδου. Είναι προφανές ότι χωρίς τους κρυμμένους κόμβους, τα απλά perceptrons ισοδυναμούν με τα γραμμικά στατιστικά μοντέλα πρόβλεψης.

Το ζήτημα του προσδιορισμού του βέλτιστου αριθμού κρυμμένων κόμβων είναι ιδιαιτέρως σημαντικό αλλά και δύσκολο. Πολλοί ερευνητές χρησιμοποιούν έναν κρυμμένο επίπεδο γιατί σύμφωνα με θεωρητικές εργασίες, το ένα επίπεδο είναι επαρκές ώστε το τεχνητό νευρωνικό δίκτυο να εκτιμήσει οποιαδήποτε πολύπλοκη μη γραμμική συνάρτηση με την επιθυμητή ακρίβεια. Το πρόβλημα όμως σε αυτήν την περίπτωση είναι ότι ένα κρυμμένο επίπεδο απαιτεί πάρα πολλούς κρυμμένους κόμβους, πράγμα που οδηγεί σε αδυναμίες στην ικανότητα γενίκευσης καθώς και στον χρόνο εκπαίδευσης. Αντίθετα, τα δίκτυα με δύο κρυμμένα επίπεδα παρουσιάζουν περισσότερα πλεονεκτήματα σε ορισμένους τύπους προβλημάτων. Έχει παρατηρηθεί ότι σε ορισμένες περιπτώσεις είχε ως αποτέλεσμα πιο συμπαγή αρχιτεκτονική η οποία απέφερε μεγαλύτερη αποτελεσματικότητα στην διαδικασία εκπαίδευσης, καθώς και μεγαλύτερη ακρίβεια αποτελεσμάτων σε σχέση με τα δίκτυα με ένα κρυμμένο επίπεδο. Ωστόσο, δεν παρατηρείται περαιτέρω βελτίωση σε δίκτυα με περισσότερα από δύο επίπεδα. Επειδή οι περισσότεροι αλγόριθμοι εκπαίδευσης είναι βαθμωτοί, το επιπλέον επίπεδο μέσω του οποίου πρέπει να διορθωθούν τα προς τα πίσω σφάλματα προκαλεί αστάθεια στην κατάβαση. Η επιτυχία οποιουδήποτε βαθμωτού αλγορίθμου βελτιστοποίησης εξαρτάται από τον βαθμό στον οποίο η κατάβαση παραμένει αμετάβλητη καθώς οι παράμετροι διαφοροποιούνται. Επίσης, όσο περισσότερα κρυμμένα επίπεδα υπάρχουν, τόσα περισσότερα τοπικά ελάχιστα θα προκύψουν. Εφόσον λοιπόν οι περισσότεροι βαθμωτοί αλγόριθμοι μπορούν να υπολογίσουν μόνο το τοπικό ελάχιστο, ο αλγόριθμος δεν είναι σε θέση να βρει το γενικό ελάχιστο, και ακόμα κι αν το έβρισκε, υπήρχε μεγάλη πιθανότητα έπειτα από πολλαπλές και χρονοβόρες επαναλήψεις το αποτέλεσμα να κολλήσει στο τοπικό ελάχιστο. Γενικά, τα δίκτυα με λιγότερους κρυμμένους κόμβους προτιμώνται καθώς παρουσιάζουν συνήθως καλύτερη ικανότητα γενίκευσης και λιγότερα προβλήματα υπερεκπαίδευσης. Αλλά τα δίκτυα με πολύ λίγους κρυμμένους κόμβους ίσως να μην έχουν την απαιτούμενη δύναμη μοντελοποίησης και μάθησης των δεδομένων. Δεν υπάρχει θεωρητική βάση για την καλύτερη επιλογή του αριθμού των κρυμμένων κόμβων. Ο συνηθέστερος τρόπος προσδιορισμού είναι μέσω πειραμάτων ή με την μέθοδο *δοκιμής-σφάλματος*. Παρόλα αυτά, έχουν προταθεί κάποιοι εμπειρικοί κανόνες, όπως ότι ο αριθμός των κρυμμένων κόμβων εξαρτάται από τον αριθμό των υποδειγμάτων εισόδου και κάθε βάρος πρέπει να έχει τουλάχιστον δέκα υποδείγματα εισόδου, καθώς και κάποιες συστηματικές προσεγγίσεις, όπως η μέθοδος αποκοπής άχρηστων κρυμμένων κόμβων, η μέθοδος πρόσθεσης κρυμμένων κόμβων κι η έρευνα πλέγματος.

3.5.1.2. Ο Αριθμός των Κόμβων Εισόδου

Δεν υπάρχει προτεινόμενος τρόπος προσδιορισμού του ακριβούς πλήθους των κόμβων εισόδου. Στις περιπτώσεις συνήθης πρόβλεψης, ο αριθμός των κόμβων εισόδου είναι συχνά προφανής και επιλέγεται σχετικά εύκολα. Στα προβλήματα πρόβλεψης χρονοσειρών (time-series), ο αριθμός των κόμβων εισόδου ανταποκρίνεται στον αριθμό των παρελθοντικών παρατηρήσεων που χρησιμοποιούνται για τη σύλληψη των υποκείμενων υποδειγμάτων στις χρονοσειρές και για την πραγματοποίηση προβλέψεων για μελλοντικές τιμές. Για καλύτερα αποτελέσματα, επιθυμούμε έναν μικρό αριθμό αναγκαίων κόμβων που μπορούν να αποκαλύψουν τα μοναδικά χαρακτηριστικά των δεδομένων. Πολύ λίγοι ή υπερβολικά πολλοί κόμβοι εισόδου ίσως να επηρεάσουν είτε την ικανότητα μάθησης είτε την ικανότητα πρόβλεψης του δικτύου. Γενικά, οι περισσότεροι μελετητές εκτελούν πειράματα για την επιλογή, ενώ άλλοι υιοθετούν εμπειρικές ιδέες.

3.5.1.3. Οι Αλληλοσυνδέσεις των Κόμβων

Οι αλληλοσυνδέσεις των κόμβων αποτελούν σημαντικό στοιχείο της αρχιτεκτονικής του δικτύου, και καθορίζει σημαντικά την συμπεριφορά του.

Ανάλογα με τον τρόπο σύνδεσής τους, τα δίκτυα μπορούν να διακριθούν ως εξής:

- πλήρως συνδεδεμένα δίκτυα. Δηλαδή, όλοι οι κόμβοι σε κάθε επίπεδο είναι πλήρως συνδεδεμένοι μόνο με τους κόμβους του επόμενου υψηλότερου επιπέδου, με μόνη εξαίρεση το επίπεδο εξόδου
- σποραδικά συνδεδεμένα δίκτυα, και
- δίκτυα με άμεσες συνδέσεις από τους κόμβους εισόδου προς τους κόμβους εξόδου. Με αυτό τον τρόπο σύνδεσης μπορούμε να πετύχουμε μεγαλύτερη ακρίβεια στην πρόβλεψη, αφού τα εν λόγω στοιχεία έχουν την δυνατότητα να χρησιμοποιηθούν για τη μοντελοποίηση των γραμμικών δομών των δεδομένων και για την ενίσχυση της δύναμης αναγνώρισης του δικτύου.

3.5.2. Η Συνάρτηση Ενεργοποίησης

Η συνάρτηση ενεργοποίησης ή αλλιώς συνάρτηση μεταφοράς, Προσδιορίζει τη σχέση μεταξύ των εισόδων και εξόδων στους κόμβους και στο δίκτυο. Παραπάνω αναφέρθηκαν οι κυριότερες συναρτήσεις ενεργοποίησης.

Παρόλο που τα περισσότερα δίκτυα χρησιμοποιούν τις ίδιες συναρτήσεις ενεργοποίησης για τους κόμβους ιδίου επιπέδου, ένα δίκτυο μπορεί να χρησιμοποιήσει διαφορετικές συναρτήσεις ενεργοποίησης για διαφορετικούς κόμβους στο ίδιο ή σε διαφορετικά επίπεδα. Στις περισσότερες περιπτώσεις, στους κρυμμένους κόμβους, η συνάρτηση ενεργοποίησης που χρησιμοποιείται είναι η logistic.

Ο κόμβος εξόδου με τις συναρτήσεις logistic ή υπερβολικής εφαπτομένης παρουσιάζει τυπικό πεδίο [0,1] ή [-1,1] αντίστοιχα. Όταν λοιπόν χρησιμοποιούνται αυτές οι μη γραμμικές συναρτήσεις στο επίπεδο εξόδου, οι επιθυμητές τιμές εξόδου πρέπει να κανονικοποιηθούν ώστε να εναρμονίζονται με την κλίμακα των πραγματικών εξόδων του δικτύου. Αν και δεν υπάρχει απόλυτη συμφωνία για το ποια συνάρτηση ενεργοποίησης πρέπει να χρησιμοποιείται στους κόμβους εξόδου, συνήθως όταν έχουμε προβλήματα ταξινόμησης, όπου οι τιμές είναι δυαδικές, χρησιμοποιείται η logistic συνάρτηση, ενώ όταν έχουμε πρόβλημα πρόβλεψης που περιλαμβάνει συνεχείς επιθυμητές τιμές χρησιμοποιείται η γραμμική συνάρτηση ενεργοποίησης. Αξιοσημείωτο είναι ότι τα feedforward νευρωνικά δίκτυα με γραμμικούς κόμβους εξόδου υπόκεινται στον περιορισμό ότι δεν μπορούν να μοντελοποιήσουν χρονοσειρές που περιέχουν τάση (trend). Συνεπώς, για αυτού του είδους νευρωνικά δίκτυα, απαιτείται προ-διαφοροποίηση ώστε να εξαλειφθούν οι τάσεις. Μέχρι σήμερα δεν έχει διερευνηθεί η συγκριτική απόδοση της χρήσης γραμμικών και μη γραμμικών συναρτήσεων ενεργοποίησης για τους κόμβους εξόδου και δεν έχουν καταγραφεί εμπειρικά αποτελέσματα για την υποστήριξη ιδιαίτερης προτίμησης σε καμία εκ των δύο.

3.5.3. Διαδικασία μάθησης

Για να χρησιμοποιηθεί ένα ΝΔ πρέπει πρώτα να εκπαιδευτεί για να μάθει. Η μάθηση συνίσταται στον προσδιορισμό των κατάλληλων συντελεστών βάρους ώστε το νευρωνικό δίκτυο να εκτελεί τους επιθυμητούς υπολογισμούς και πραγματοποιείται με τη βοήθεια αλγορίθμων που είναι γνωστοί ως κανόνες μάθησης. Ο ρόλος των συντελεστών βάρους μπορεί να ερμηνευτεί ως αποθήκευση γνώσης η οποία παρέχεται μέσω παραδειγμάτων. Με αυτόν τον τρόπο τα νευρωνικό δίκτυο μαθαίνουν το περιβάλλον τους δηλαδή το φυσικό μοντέλο που παρέχει τα δεδομένα και βελτιώνουν τη συμπεριφορά τους στο πέρασμα του χρόνου.

Έτσι ένα νευρωνικό δίκτυο μαθαίνει για το περιβάλλον του μέσω μιας επαναληπτικής διαδικασίας ανανέωσης των συναπτικών βαρών και κατωφλίων. Κάθε αλγόριθμος μάθησης προσφέρει έναν άλλο τρόπο προσαρμογής των συναπτικών βαρών. Τα προβλήματα μάθησης τα οποία επιλύουν οι αντίστοιχοι αλγόριθμοι εξαρτώνται και από το περιβάλλον στο οποίο εργάζεται κάθε νευρωνικό δίκτυο. Έτσι διαφορετικά μοντέλα του περιβάλλοντος οδηγούν σε διαφορετικά μοντέλα εκπαίδευσης.

<u>Υπάρχουν 3 βασικά είδη μάθησης:</u>

- Μάθηση υπό επίβλεψη (supervised learning)
    - Ενισχυτική μάθηση
    - Ανταγωνιστική μάθηση
    - Τυχαία μάθηση

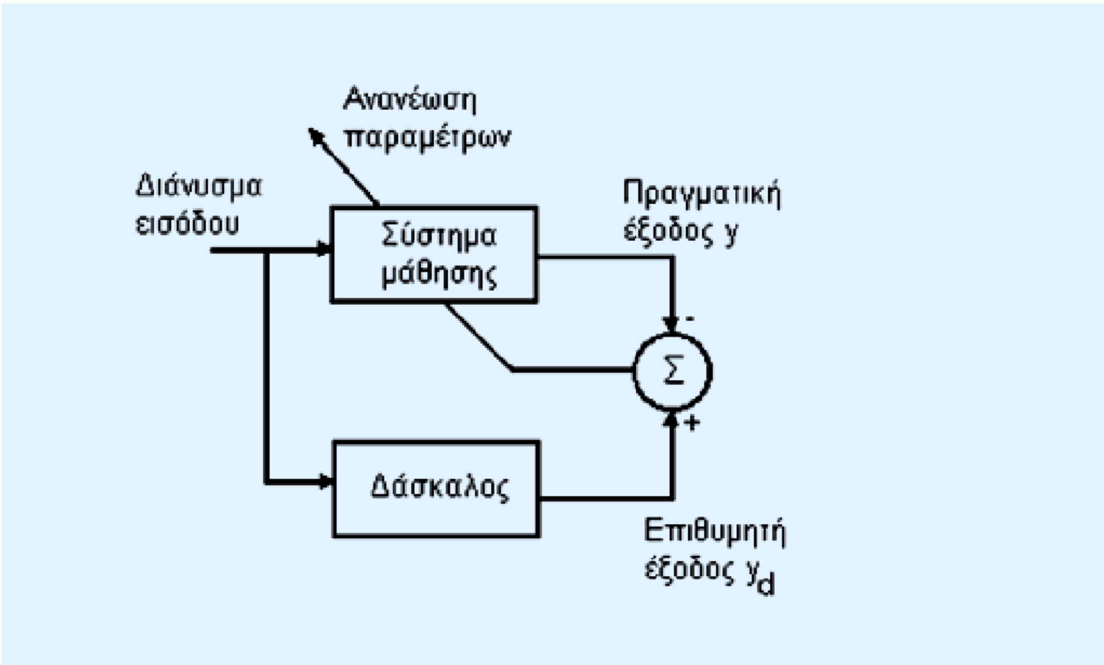
- Μάθηση χωρίς επίβλεψη (unsupervised learning)
- Βαθμολογημένη μάθηση (graded learning)

3.5.3.1. Μάθηση υπό επίβλεψη

Όπως μπορεί να παρατηρηθεί από το παρακάτω σχήμα αναπαράστασης της δομής ενός δικτύου υπό επίβλεψη, η δομή αποτελείται από δύο βασικές συνιστώσες, το σύστημα μάθησης και από τον δάσκαλο. Ο δάσκαλος περιλαμβάνει τη γνώση που απαιτείται ώστε να διδάξει στο σύστημα μάθησης τις επιθυμητές εξόδους για το σύνολο εισόδων εκπαίδευσης. Όταν ο δάσκαλος και το νευρωνικό δίκτυο λαμβάνουν ένα διάνυσμα εισόδου εκπαίδευσης, ο δάσκαλος δίνει στο νευρωνικό δίκτυο μια επιθυμητή έξοδο η οποία αναπαριστά την βέλτιστη δράση που πρέπει να εμφανίζει το νευρωνικό δίκτυο. Οι παράμετροι του νευρωνικού δικτύου ανανεώνονται βάσει του διανύσματος εκπαίδευσης και του διανύσματος σφάλματος (δηλαδή της διαφοράς πραγματικής και επιθυμητής απόκρισης του δικτύου). Η συνάρτηση κόστους ορίζεται ως:

$$J(w)=(1/2)*\ E[(y(t)-y_d(t)^2]$$

όπου w είναι το διάνυσμα των προς επιλογή ελεύθερων παραμέτρων του συστήματος μάθησης (δηλαδή του νευρωνικού δικτύου), η ανανέωση των παραμέτρων παίρνει την μορφή του αλγορίθμου διόρθωσης σφάλματος. Η σταδιακή ανανέωση των παραμέτρων κάνει τελικά το νευρωνικό δίκτυο μάθησης να μιμείται την δεδομένη επιθυμητή συμπεριφορά.



<u>Σχήμα 13: Δίκτυο υπό επίβλεψη</u>

3.5.3.1.1. Η ενισχυτική μάθηση

Η ενισχυτική μάθηση (reinforcement learning) είναι μια ειδική περίπτωση της εποπτευόμενης μάθησης, όπου το εξωτερικό περιβάλλον παρέχει μόνο την πληροφορία ότι η έξοδος του δικτύου είναι καλή ή κακή, αντί να δίνει τη σωστή έξοδο. Αποτιμάται δηλαδή η αποτελεσματικότητα του συστήματος.

Το σήμα μπορεί να είναι συνεχές, διακριτό στο [0,1] και απεικονίζει την απόκριση του περιβάλλοντος στην εκτέλεση μιας ενέργειας με την μορφή ενός βαθμού επιτυχίας ή αποτυχίας. Ένα σήμα επιτυχίας έχει ως αποτέλεσμα την ακόμη μεγαλύτερη ενίσχυση της ενέργειας που επιλέχτηκε ενώ ένα σήμα αποτυχίας αποτρέπει την επανάληψη της επιλογής ενέργειας όταν ξαναεμφανιστεί το ίδιο σήμα εισόδου. Στην περίπτωση της ενισχυτικής μάθησης λέγεται ότι το εξωτερικό περιβάλλον συμπεριφέρεται ως ‘’κριτής’’ (critic).

Η ενισχυτική μάθηση θα μπορούσε να αποτελεί μια τρίτη κατηγορία μάθησης από μόνη της και να μην την αντιμετωπίζουν ως μια υποπερίπτωση της εποπτευόμενης μάθησης.

3.5.3.1.2. Ανταγωνιστική μάθηση

Στην ανταγωνιστική μάθηση οι τεχνητοί νευρώνες ανταγωνίζονται, κατά κάποιο τρόπο, μεταξύ τους και μόνο αυτός με τη μεγαλύτερη απόκριση σε δοθείσα είσοδο τροποποιεί τα βάρη του. Έτσι έχουμε ένα ανταγωνισμό στους νευρώνες εξόδου του δικτύου με αποτέλεσμα σε κάθε χρονική στιγμή να είναι ενεργός ένας μόνο νευρώνας.

3.5.3.1.3 Τυχαία μάθηση

Στην τυχαία μάθηση οι μεταβολές στα βάρη εισάγονται τυχαία και ανάλογα με το αν η έξοδος βελτιώνεται ή όχι με βάση κάποια προκαθορισμένα από το χρήστη κριτήρια, οι μεταβολές αυτές υιοθετούνται ή απορρίπτονται.

3.5.3.2. Μάθηση χωρίς επίβλεψη

Αυτός ο τύπος μάθησης, που καλείται αυτοοργανούμενη μάθηση, χρησιμοποιεί μόνο τα διανύσματα εισόδου για την υλοποίηση της εκπαίδευσης. Ένα σύστημα μη επιβλεπόμενης μάθησης εξελίσσεται με τέτοιο τρόπο ώστε να εξάγει χαρακτηριστικά ή κανονικότητες από τα παρουσιαζόμενα πρότυπα, χωρίς ωστόσο να έχει την πληροφορία για το ποιες έξοδοι ή ποιες κατηγορίες συσχετίζονται με τα χαρακτηριστικά εισόδου. Με άλλα λόγια το σύστημα μάθησης εντοπίζει ή κατηγοριοποιεί τα διανύσματα εισόδου χωρίς καμία εκ των προτέρων πληροφόρηση από το περιβάλλον. Εξ’ αιτίας αυτών, η μη επιβλεπόμενη μάθηση συχνά χρησιμοποιείται σε προβλήματα ομαδοποίησης, εξαγωγής εσωτερικών χαρακτηριστικών και ανίχνευσης συμμετριών. Με αυτήν τη μέθοδο το νευρωνικό δίκτυο αναπτύσσει συγκεκριμένες εσωτερικές αναπαραστάσεις οι οποίες κωδικοποιούν τα διάφορα διανύσματα εισόδου.

3.5.3.3. Βαθμολογημένη μάθηση

Η έξοδος χαρακτηρίζεται ως "καλή" ή "κακή" με βάση μια αριθμητική κλίμακα και τα βάρη αναπροσαρμόζονται με βάση αυτό το χαρακτηρισμό.

3.5.4. Δεδομένα εκπαίδευσης και δοκιμής

Ένα ακόμα σημαντικό βήμα στην κατασκευή των νευρωνικών δικτύων, που επιδρά στην απόδοση και στην αξιολόγηση του δικτύου, είναι ο ορισμός των δεδομένων εκπαίδευσης (insample data) και των δεδομένων δοκιμής (out-of-sample ή hold-out sample). Το δείγμα εκπαίδευσης χρησιμοποιείται για την ανάπτυξη του μοντέλου ΤΝΔ και το δείγμα δοκιμής για την εκτίμηση της ικανότητας πρόβλεψης του μοντέλου. Ορισμένες φορές χρησιμοποιείται και ένα τρίτο δείγμα, το δείγμα αξιολόγησης (validation sample), ώστε να αποφευχθεί το πρόβλημα υπερεκπαίδευσης ή να προσδιοριστεί το σημείο τερματισμού της διαδικασίας εκπαίδευσης.

Η διαδικασία εκπαίδευσης έχει συνήθως ως εξής:

- Πρώτον, παραδείγματα των συνόλων εκπαίδευσης εισάγονται στους κόμβους εισόδου, οι τιμές ενεργοποίησης των κόμβων εισόδου σταθμίζονται και συσσωρεύονται σε κάθε κόμβο στο πρώτο κρυμμένο επίπεδο

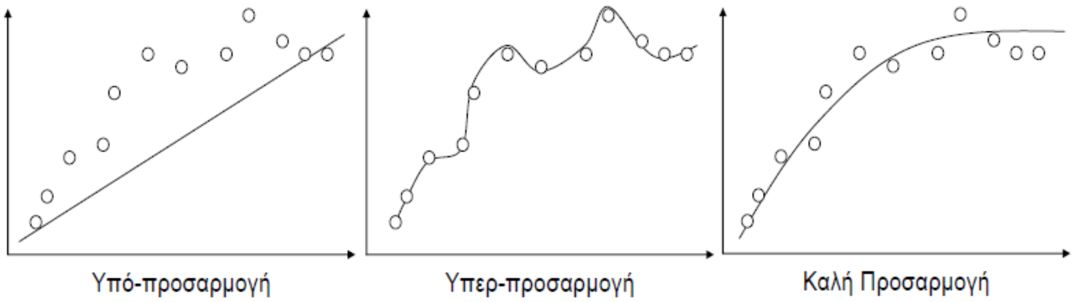
- το σύνολο μορφοποιείται από τη συνάρτηση ενεργοποίησης στις τιμές ενεργοποίησης των κόμβων

- διαδοχικά αναπτύσσεται σε μια είσοδο μέσα στους κόμβους του επόμενου επιπέδου, έως ότου τελικά βρεθούν οι τιμές ενεργοποίησης των εξόδων.

Ο αλγόριθμος εκπαίδευσης χρησιμοποιείται για τον προσδιορισμό των βαρών που ελαχιστοποιούν ορισμένα κριτήρια ολικού σφάλματος όπως το άθροισμα των τετραγωνικών σφαλμάτων (sum of squares errors SSE), ή των μέσων τετραγωνικών σφαλμάτων (mean squared errors MSE). Συνεπώς, η εκπαίδευση του δικτύου είναι στην πράξη ένα μη γραμμικό πρόβλημα ελαχιστοποίησης χωρίς περιορισμούς. Τα τεχνητά νευρωνικά δίκτυα είναι προγραμματισμένα να μαθαίνουν έως ένα δεδομένο όριο σφάλματος. Σε ένα καλά εκπαιδευμένο νευρωνικό δίκτυο το όριο σφάλματος και το πραγματικό σφάλμα πρέπει να ταυτίζονται. Παρόλο που δεν υπάρχει γενικός κανόνας λύσης στο εν λόγω πρόβλημα, πρέπει να ληφθούν υπόψη, στην απόφαση, αρκετοί παράγοντες, όπως τα χαρακτηριστικά του προβλήματος, καθώς και το είδος και το μέγεθος των δεδομένων. Είναι αποφασιστικής σημασίας τόσο το σύνολο εκπαίδευσης όσο και το σύνολο δοκιμής να είναι αντιπροσωπευτικά του δείγματος ή του υποκείμενου μηχανισμού. Ακατάλληλος διαχωρισμός των συνόλων εκπαίδευσης και δοκιμής θα επηρεάσει την επιλογή της βέλτιστης δομής ΤΝΔ και την εκτίμηση της προγνωστικής απόδοσης του ΤΝΔ. Οι έρευνες που έχουν διεκπεραιωθεί προσφέρουν ελάχιστη καθοδήγηση στην επιλογή των δειγμάτων εκπαίδευσης και δοκιμής.

Οι περισσότεροι μελετητές επιλέγουν βάσει των κανόνων 90% έναντι 10%, 80% έναντι 20% ή 70% έναντι 30% κτλ. Ορισμένοι επιλέγουν βάσει του προβλήματος που μελετάται. Ένας ακόμη σχετικός παράγοντας είναι το μέγεθος του δείγματος. Δεν υπάρχει συγκεκριμένος κανόνας για τις απαιτήσεις του μεγέθους δείγματος για κάθε πρόβλημα. Η ποσότητα των δεδομένων για την εκπαίδευση του δικτύου εξαρτάται από τη δομή του δικτύου, τη μέθοδο εκπαίδευσης και την πολυπλοκότητα του εκάστοτε προβλήματος ή την ποσότητα του θορύβου στα διαθέσιμα δεδομένα. Γενικά, όπως σε κάθε στατιστική προσέγγιση, το μέγεθος δείγματος σχετίζεται στενά με την απαιτούμενη ακρίβεια του προβλήματος. Όσο μεγαλύτερο το μέγεθος δείγματος, τόσο πιο ακριβή θα είναι τα αποτελέσματα. Έχει καταγραφεί σε ερευνητικές μελέτες ότι καθώς αυξάνεται το μέγεθος, τα ΤΝΔ πρόβλεψης αποδίδουν καλύτερα.
Εξαιρετικά σημαντικός παράγοντας στην εκπαίδευση ενός νευρωνικού δικτύου είναι ο αριθμός εποχών εκπαίδευσης. Ως εποχή (epoch) εννοούμε την χρήση όλων των

δεδομένων εκπαίδευσης (από το δίκτυο) για μία φορά. Μικρός αριθμός εποχών εκπαίδευσης συνήθως οδηγεί σε άσχημα αποτελέσματα ενώ αντίθετα μεγάλος αριθμός εποχών εκπαίδευσης οδηγεί στην υπερεκπαίδευση και απαιτεί μεγάλο υπολογιστικό χρόνο. Υπερεκπαίδευση έχουμε όταν το νευρωνικό δίκτυο προσαρμοστεί πολύ καλά στα δεδομένα εκπαίδευσης και μόνο σε αυτά. Οι αναφορές μικρός και μεγάλος αριθμός εποχών είναι σχετικές και αναφέρονται στο εκάστοτε πρόβλημα.



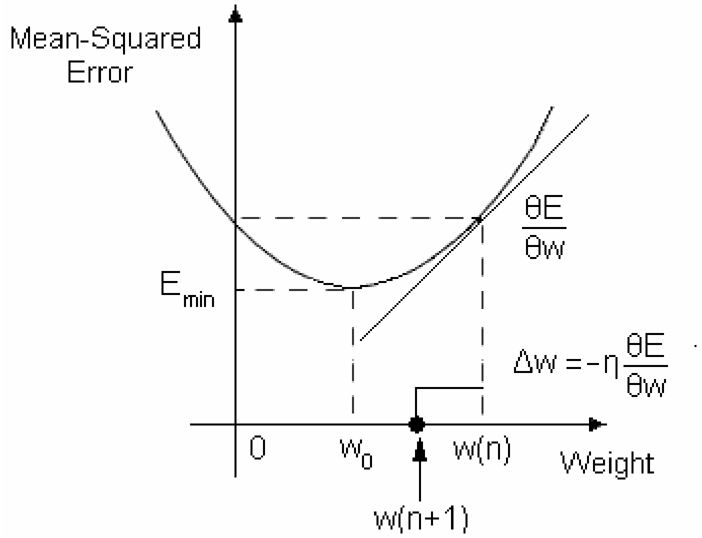
<u>Εικόνα 3</u>

3.6. Αλγόριθμοι εκπαίδευσης

Στόχος των αλγορίθμων μάθησης είναι η συστηματική και ταχεία εύρεση των συναπτικών βαρών του δικτύου ώστε να ελαχιστοποιείται κάποιο μέτρο του σφάλματος μεταξύ των πραγματικών και των επιθυμητών εξόδων του δικτύου. Η δυνατότητα εφαρμογής των διαφόρων αλγορίθμων μάθησης που έχουν αναπτυχθεί, εξαρτάται από τον τύπο του ΤΝΔ και από την κατηγοριοποίηση των διαδικασιών μάθησης.

3.6.1. Ο Κανόνας Μάθησης Δέλτα

Ο κανόνας Δέλτα θεωρούμε ότι είναι μια διαδικασία ανίχνευσης κατηφορικής κλίσης *(gradient-descent search procedure)* η οποία εκτελείται σε κάθε επανάληψη. Σύμφωνα με αυτήν, η αλλαγή του βάρους γίνεται σε κατεύθυνση αντίθετη με αυτήν που έχει η παράγωγος (ως προς το βάρος) της επιφάνειας του σφάλματος, έχοντας σαν αποτέλεσμα την απότομη κατάβαση σφάλματος προς το ελάχιστο σημείο του όπως φαίνεται και στο παρακάτω Σχήμα.



Σχήμα 14

Το σφάλμα μεταξύ της πραγματικής και της επιθυμητής τιμής ενός νευρώνα είναι:

$$e = d - f(\sigma) = d - f(\langle w, x \rangle)$$

Η ευαισθησία του σφάλματος e ορίζεται από τη παρακάτω μερική παράγωγο:

$$\frac{\partial e}{\partial w} = -\frac{\partial y}{\partial w} = -\frac{\partial f(\sigma)}{\partial w} = -\frac{\partial f(\sigma)\partial \sigma}{\partial \sigma \partial w} = -f'(\sigma) \cdot x = -g(\sigma) \cdot x$$

Η οποία υπολογίζεται με τον αλυσιδωτό κανόνα παραγώγισης. Για να ορίζεται η ευαισθησία παντού πρέπει να ισχύει η g(σ)=f'(σ) για κάθε σ. Δηλαδή, η σχέση εισόδου-εξόδου του μη γραμμικού νευρώνα πρέπει να είναι παραγωγίσιμη.

Ακολουθώντας τα ίδια βήματα, προσεγγίζουμε τη μερική παράγωγο του τετραγώνου του σφάλματος:

$$\hat{\gamma} = \frac{\partial e^2}{\partial w} = 2e\frac{\partial e}{\partial w} = -2e\frac{\partial f}{\partial w} = -2e\frac{\partial f \partial \sigma}{\partial \sigma \partial w} = -2egx$$

Οπότε το w θα δίνεται από την παρακάτω επαναληπτική σχέση:

$$w^{m+1} = w^m - \mu\gamma^m = w^m + \Delta w^m = w^m + 2me^m g^m x^m$$

όπου m ο δείκτης επανάληψης.

Η επιλογή του συντελεστή **μ** είναι αυθαίρετη. Μικρή τιμή του συντελεστή **μ** οδηγεί σε αργή σύγκλιση, ενώ μεγάλη τιμή του συντελεστή **μ** μπορεί να προκαλέσει απόκλιση του αλγορίθμου.

Υπάρχουν πολλές τεχνικές για την αύξηση της ταχύτητας σύγκλισης του αλγόριθμου εκμάθησης, με την αυτόματη μεταβολή του συντελεστή **μ**, ανάλογα με την πορεία της μάθησης. Ένας απλός τρόπος είναι το να ξεκινήσει η μάθηση λαμβάνοντας τυχαίες τιμές και για τα βάρη του συστήματος και ένα συντηρητικό συντελεστή **μ** . Μετά της πρώτες επαναλήψεις του αλγόριθμου εξετάζεται η εξέλιξη της νόρμας του μεγέθους των διορθώσεων στα βάρη του συστήματος $\|\Delta w_m\|$. Αν ο ρυθμός εξέλιξης της νόρμας είναι μικρότερος από ένα προκαθορισμένο όριο **ε** τότε η τιμή του συντελεστή μ διπλασιάζεται, ενώ αν ο ρυθμός εξέλιξης είναι μεγαλύτερος από ένα, επίσης προκαθορισμένο, όριο **δ** τότε η τιμή του συντελεστή **μ** μειώνεται στο μισό. Με τον παραπάνω απλό τρόπο επιτυγχάνεται μεταβαλλόμενος ρυθμός μάθησης.

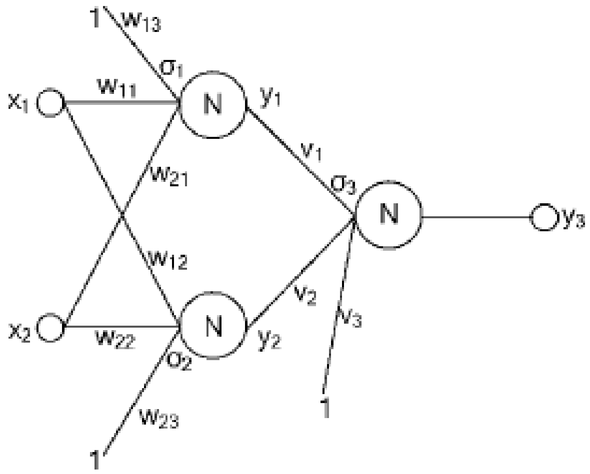
3.6.2. Ο Αλγόριθμος Μάθησης Με Οπισθόδρομη Διάδοση(backpropagation algorithm)

Έστω ένα δίκτυο δύο στρωμάτων με δύο νευρώνες στο στρώμα εισόδου και ένα νευρώνα στο στρώμα εξόδου (το δίκτυο παρουσιάζεται στο σχήμα που ακολουθεί). Η γενίκευση του αλγορίθμου είναι σχετικά εύκολη, αλλά την δυσκολεύει το γεγονός του ότι απαιτούνται πολλαπλοί δείκτες στα συναπτικά βάρη τα οποία καθιστούν τις εξισώσεις ιδιαίτερα δύσχρηστες. Από το σχήμα είναι προφανές ότι τα σήματα στις εισόδους των νευρώνων του πρώτου στρώματος είναι:

$\sigma 1 = w11x1 + w21x2 + w13 \quad y1 = f(\sigma 1)$
$\sigma 2 = w12x1 + w22x2 + w23 \quad y2 = f(\sigma 2)$

Ομοίως η σταθμισμένη έξοδος του δευτέρου στρώματος είναι:

$\sigma 3 = v1y1 + v2y2 + v3 \quad y3 = f(\sigma 3)$



Σχήμα 15: Ένα απλό ΤΝΔ με 2 εισόδους και 1 έξοδο (2-1)

Το σφάλμα του δικτύου είναι η διαφορά μεταξύ της πραγματικής και της επιθυμητής εξόδου του δικτύου, δηλαδή:

e = d- y3 .

Σε κάθε επανάληψη το σφάλμα το οποίο υπολογίζεται είναι :

e = d – f(y3) .

Άρα απαιτείται η μερική παράγωγος του τετραγώνου του σφάλματος για το κάθε βάρος. Για τα βάρη του πρώτου και του δεύτερου στρώματος η μερικές παράγωγοι του τετραγώνου του σφάλματος είναι :

$$\frac{\partial e^2}{\partial(w_{ij})} = 2e\frac{\partial e}{\partial(v_i)} \quad , \ i = 1{,}2 \ \kappa\alpha\iota \ j = 1{,}3$$

$$\frac{\partial e^2}{\partial(v_i)} = 2e\frac{\partial e}{\partial(v_i)} \quad , \ i = 1{,}2 \ \kappa\alpha\iota \ j = 1{,}3$$

Αν ακολουθήσουμε τα βήματα τα οποία έχουν παρουσιαστεί παραπάνω μπορούμε να εκφράσουμε τον επαναληπτικό αλγόριθμο μάθησης:

$$w^{m+1} = w^m + \Delta w^m = w^m + 2\mu e^m \frac{\partial e^m}{\partial w^m}$$

και

$$v^{m+1} = v^m + \Delta v^m = w^m + 2\mu e^m \frac{\partial e^m}{\partial v^m}$$

Ο αλγόριθμος απαιτεί τον υπολογισμό των ευαισθησιών του σφάλματος σε σχέση με τα βάρη των διασυνδέσεων των νευρώνων , δηλαδή τις παρακάτω μερικές παραγώγους:

$$\frac{\partial e}{\partial w}$$

και

$$\frac{\partial e}{\partial v}$$

Οι πρώτες τρείς μερικές παράγωγοι , δηλαδή οι παράγωγοι της ευαισθησίας του σφάλματος στα βάρη των νευρώνων, θα βρεθούν με χρήση του αλυσιδωτού κανόνα παραγώγισης και είναι:

$$\frac{\partial e}{\partial w_{11}} = -\frac{\partial y}{\partial w_{11}} = \frac{\partial f(\sigma_3)\partial \sigma_3}{\partial \sigma_3 \partial w_{11}} = -\frac{\partial f(\sigma_3)\partial f(\sigma_1)\partial \sigma_1}{\partial \sigma_3 \partial f(\sigma_1)\partial \sigma_1 \partial w_{11}} = -g(\sigma_3)g(\sigma_1)v_1 x_1$$

$$\frac{\partial e}{\partial w_{12}} = -\frac{\partial y}{\partial w_{12}} = \frac{\partial f(\sigma_3)\partial \sigma_3}{\partial \sigma_3 \partial w_{12}} = -\frac{\partial f(\sigma_3)\partial f(\sigma_1)\partial \sigma_1}{\partial \sigma_3 \partial f(\sigma_1)\partial \sigma_1 \partial w_{12}} = -g(\sigma_3)g(\sigma_1)v_1 x_2$$

$$\frac{\partial e}{\partial w_{13}} = -\frac{\partial y}{\partial w_{13}} = \frac{\partial f(\sigma_3)\partial \sigma_3}{\partial \sigma_3 \partial w_{13}} = -\frac{\partial f(\sigma_3)\partial f(\sigma_1)\partial \sigma_1}{\partial \sigma_3 \partial f(\sigma_1)\partial \sigma_1 \partial w_{13}} = -g(\sigma_3)g(\sigma_1)v_1 x_3$$

Αντίστοιχα υπολογίζονται και οι υπόλοιπες μερικές παράγωγοι, οι οποίες αφορούν τα βάρη που αντιστοιχούν στη δεύτερη είσοδο. Αξίζει να σημειωθεί ότι οι αναλυτικές σχέσεις των μη-γραμμικών συναρτήσεων και των παραγώγων τους είναι γνωστές και άρα ο υπολογισμός τους γίνεται χωρίς ιδιαίτερη δυσκολία.

Στην περίπτωση που το μη-γραμμικό στοιχείο του νευρώνα είναι σιγμοειδές τότε η παράγωγος είναι:

$$f(\sigma) = \frac{1 - e^{-\sigma}}{1 + e^{-\sigma}} \quad \kappa\alpha\iota \quad g(\sigma) = f'(\sigma) = \frac{2\sigma}{(1 + e^{-\sigma})^2}$$

Με τον ίδιο τρόπο, υπολογίζονται οι ευαισθησίες στα βάρη του δεύτερου στρώματος.

$$\frac{\partial e}{\partial v_1} = -\frac{\partial y_3}{\partial v_1} = -\frac{\partial f(\sigma_3)\,\partial \sigma_3}{\partial \sigma_3\,\partial v_1} = -g(\sigma_3)y_1$$

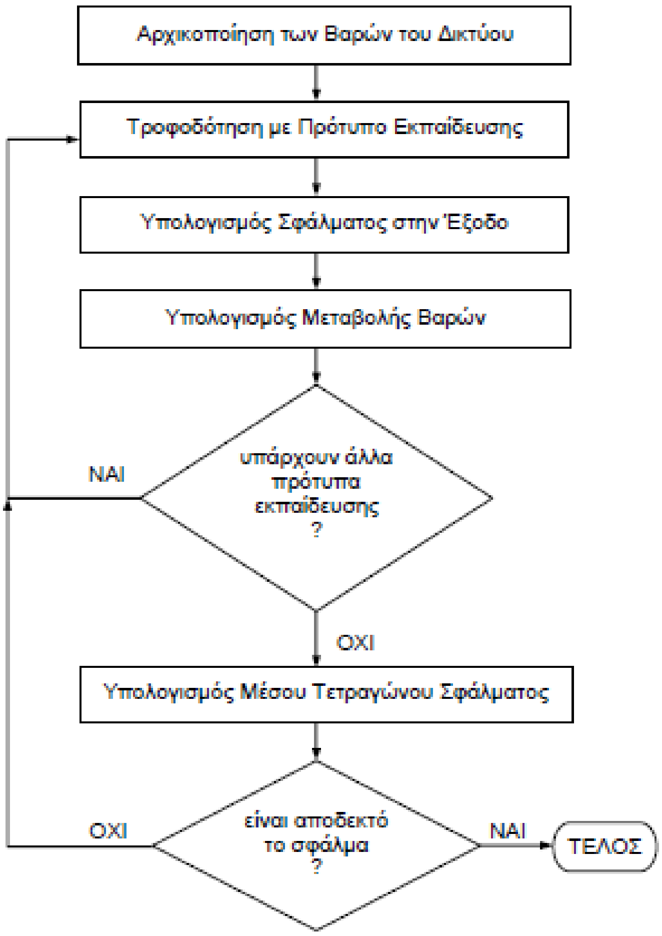
$$\frac{\partial e}{\partial v_2} = -\frac{\partial y_3}{\partial v_2} = -\frac{\partial f(\sigma_3)\,\partial \sigma_3}{\partial \sigma_3\,\partial v_2} = -g(\sigma_3)y_2$$

$$\frac{\partial e}{\partial v_3} = -\frac{\partial y_3}{\partial v_3} = -\frac{\partial f(\sigma_3)\,\partial \sigma_3}{\partial \sigma_3\,\partial v_3} = -g(\sigma_3)y_3$$

Οι παραπάνω εξισώσεις είναι απλές εφόσον διασχίζεται μόνο ένα στρώμα. Ο αλγόριθμος οπισθόδρομης διάδοσης παίρνει το όνομα του από το γεγονός ότι για να υπολογιστούν τα βάρη γίνεται ανάκληση στη ροή των υπολογισμών. Έτσι με κάποιες τυχαίες αρχικές τιμές για τα βάρη των στρωμάτων και με δεδομένες τις τιμές των εισόδων, γίνεται η πρώτη υπολογιστική διαδρομή από την οποία υπολογίζονται οι μεταβλητές σ_1, σ_2, σ_3 και συνεπώς y_1 y_2, y_3. Με τη γνώση αυτή υπολογίζονται οι διάφορες συναρτήσεις των μη-γραμμικών σχέσεων εισόδων-εξόδου και οι παράγωγοι τους (f,g). Σε κάθε επανάληψη του αλγορίθμου υπολογίζονται οι μεταβολές ή διορθώσεις ΔW_m στα βάρη του στρώματος της εισόδου και Δv_m στα βάρη του στρώματος της εξόδου. Οι μεταβολές αυτές προστίθενται στις προηγούμενες τιμές των βαρών για να επαναληφθεί η διαδικασία μέχρι να επιτευχθεί το μέτρο του σφάλματος οπότε και τερματίζεται η διαδικασία μάθησης.

Για σύνθετα ΤΝΔ απαιτούνται δεκάδες χιλιάδες επαναλήψεις για τη μάθηση των βαρών οπότε η διαδικασία μάθησης είναι αρκετά χρονοβόρα και απαιτεί ταχύτατους υπολογιστές. Αξίζει να σημειωθεί ότι ο συγκεκριμένος αλγόριθμος εποπτευόμενης μάθησης αν και είναι πολύ διαδεδομένος παρουσιάζει αργή σύγκλιση, οπότε η χρήση του γίνεται προβληματική για μεγάλα δίκτυα. Μια παραλλαγή του παραπάνω αλγόριθμου είναι ο αλγόριθμος μάθησης με ορμή (momentum) ο οποίος είναι ιδιαίτερα αποτελεσματικός στη μάθηση νευρωνικών ελεγκτών. Στην περίπτωση αυτή προστίθεται στη εξίσωση αναπροσαρμογής των βαρών ένας ακόμη όρος που είναι ανάλογος της μεταβολής των βαρών με αποτέλεσμα τη σημαντική επιτάχυνση τουαλγορίθμου.

Παρακάτω παρατίθεται η διαδικασία εκπαίδευσης με τη μορφή σχεδιαγράμματος. Η όλη διαδικασία βασίζεται στην ανάστροφη μετάδοση του σφάλματος (back propagation):



Σχήμα 16: Διαδικασία εκπαίδευσης ενός ΤΝΔ

3.7. Βασικές ιδιότητες των τεχνητών νευρωνικών δικτύων

- Έχουν την ικανότητα να μαθαίνουν μέσω παραδειγμάτων και να συλλαμβάνουν περίπλοκες λειτουργικές σχέσεις μεταξύ των δεδομένων. Συνεπώς τα ΤΝΔ είναι κατάλληλα για προβλήματα των οποίων οι λύσεις απαιτούν γνώσεις δύσκολα καθορισμένες για τις οποίες τα δεδομένα ή οι παρατηρήσεις είναι ανεπαρκή. Υπό αυτή την έννοια συγκαταλέγονται στις πλέον μη γραμμικές, μη παραμετρικές με πολλαπλές μεταβλητές στατιστικές μεθόδους.

- Υπάρχει η δυνατότητα θεώρησής τους ως κατανεμημένη μνήμη (distributed memory) και ως μνήμη συσχέτισης (associative memory).

- Έχουν μεγάλη ανοχή σε σφάλματα (fault- tolerant).

- Έχουν την ικανότητα να αναγνωρίζουν πρότυπα (pattern recognition).

- Έχουν τη δυνατότητα γενίκευσης. Έπειτα από τη διαδικασία μάθησης των δεδομένων, τα ΤΝΔ έχουν συχνά τη δυνατότητα εξαγωγής τεκμηριωμένων συμπερασμάτων των μη ορατών μερών του δείγματος ακόμη και εάν το δείγμα περιέχει πληροφορίες με θόρυβο.

- Η πρόβλεψη πραγματοποιείται, μέσω προγνωστικών της μελλοντικής συμπεριφοράς (το μη ορατό μέρος) από παραδείγματα παρελθοντικής συμπεριφοράς.

- Τα τεχνητά νευρωνικά δίκτυα, που αποτελούν μη γραμμικές, οδηγούμενες από τα δεδομένα, προσεγγίσεις σε αντίθεση μη γραμμικά μοντέλα, είναι ικανά για την εκτέλεση μη γραμμικής μοντελοποίησης χωρίς να απαιτούν a priori γνώση σχετικά με τη σχέση των μεταβλητών εισόδου και εξόδου. Συνεπώς συνιστούν ένα πιο γενικό και ευέλικτο εργαλείο μοντελοποίησης για την διεξαγωγή προβλέψεων.

3.8 . Εφαρμογές νευρωνικών δικτύων

Τα νευρωνικά δίκτυα και τα νεύρο-ασαφή συστήματα χρησιμοποιούνται σε μεγάλο αριθμό εφαρμογών. Πιο συγκεκριμένα , οι κυριότεροι τομείς χρήσεις τους είναι οι εξής:

- **Αεροπορία:** Υψηλής απόδοσης αυτόματοι πιλότοι αεροπλάνων, προσομοιωτές πτήσης, συστήματα αυτομάτου ελέγχου αεροπλάνων, συστήματα ανίχνευσης βλαβών.
- **Αυτοκίνηση:** Αυτοκινούμενα συστήματα αυτόματης πλοήγησης
- **Άμυνα:** Πλοήγηση όπλων, ανίχνευση στόχων, νέα είδη αισθητήρων, sonar, radar, ψηφιακή επεξεργασία σημάτων, συμπίεση δεδομένων, εξαγωγή χαρακτηριστικών, αναγνώριση σήματος / εικόνας.
- **Ηλεκτρονική:** Πρόβλεψη ακολουθίας κωδίκων, μορφοποίηση ολοκληρωμένων κυκλωμάτων, έλεγχος διεργασιών, διάγνωση βλαβών ολοκληρωμένων κυκλωμάτων, μηχανική όραση, σύνθεση φωνής.
- **Κοινωνική ασφάλιση:** Αξιολόγηση εφαρμοζόμενης πολιτικής, βελτιστοποίηση παραγωγής.
- **Τραπεζικές εφαρμογές:** Αναγνώστες επιταγών και άλλων παραστατικών, συστήματα αξιολόγησης αιτήσεων δανειοδότησης . Μια ποικιλία από κομμάτια πληροφοριών, τα οποία είναι συνήθως γνωστά για ένα απαιτούμενο

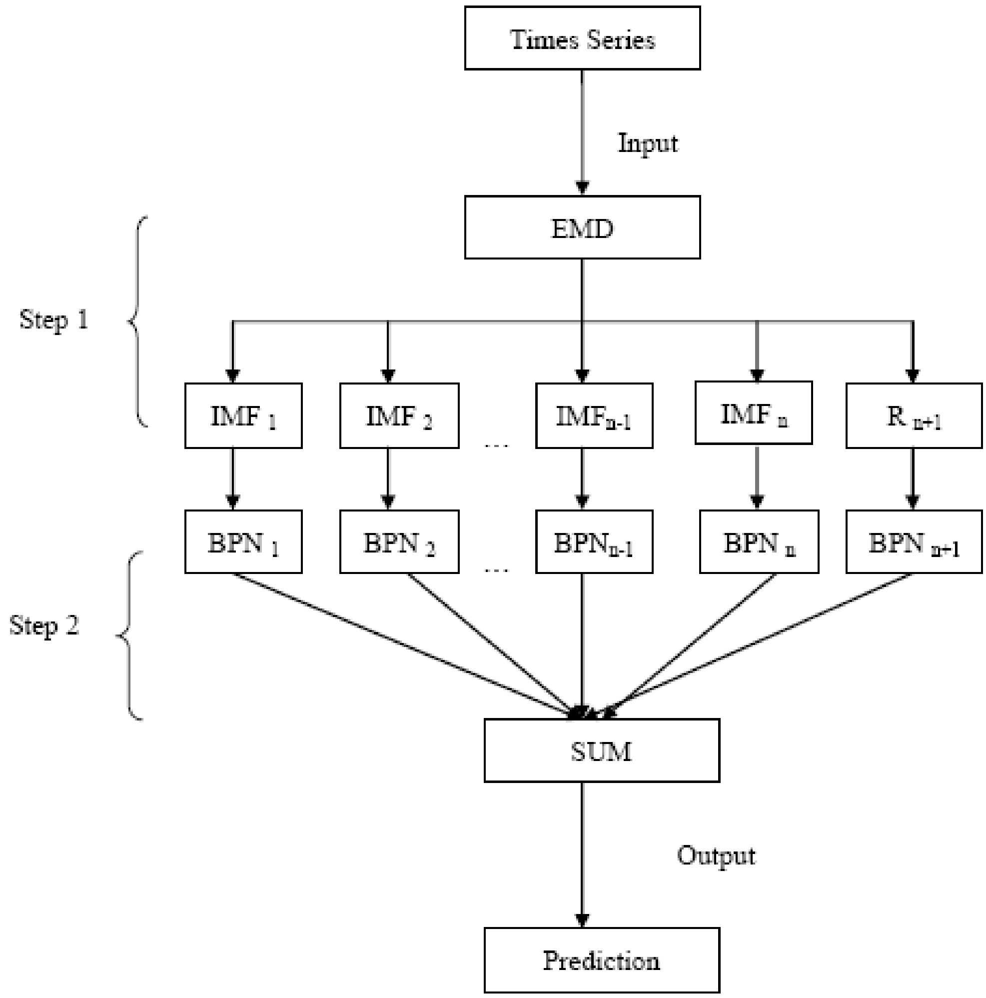
δάνειο. Για παράδειγμα, η ηλικία του αιτούντος, η εκπαίδευση, το επάγγελμα και πολλά άλλα στοιχεία που μπορεί να είναι διαθέσιμα. Μετά την εκπαίδευση ενός νευρωνικού δικτύου σε ιστορικά δεδομένα η ανάλυση μπορεί να εκτοπίσει τα πιο κατάλληλα και σχετικά χαρακτηριστικά και να τα χρησιμοποιήσει για την ταξινόμηση των αιτούντων ως χαμηλού ή υψηλού κινδύνου.

- **Βιομηχανία:** Βιομηχανικός έλεγχος διεργασιών, ανάλυση και σχεδίαση προϊόντων, συστήματα ποιοτικού ελέγχου, διάγνωση βλαβών διεργασιών και μηχανών, ανάλυση σχεδιασμού χημικών προϊόντων, δυναμική μοντελοποίηση συστημάτων χημικών διεργασιών, σχεδιασμός και διοίκηση. Τα νευρωνικά δίκτυα μπορούν να συμβάλλουν στη μείωση του κόστους με την εξασφάλιση της πρόσθετης εμπειρογνωμοσύνης για τον προγραμματισμό προληπτικής συντήρησης των μηχανημάτων. Ένα νευρωνικό δίκτυο, λοιπόν, μπορεί να εκπαιδευτεί με τέτοιο τρόπο, ώστε να διακρίνει από τους ήχους τους οποίους παράγει μια μηχανή είτε αν εκτελεί κανονικά τις λειτουργίες της, είτε βρίσκεται στα πρόθυρα εμφάνισης οποιασδήποτε δυσλειτουργίας. Μετά από αυτήν την περίοδο εκπαιδευτικής κατάρτισης, η εμπειρία του ίδιου δικτύου είναι δυνατό να χρησιμοποιηθεί με σκοπό την προειδοποίηση ενός τεχνικού για κάποια επικείμενη βλάβη προτού συμβεί και ενδεχομένως προκαλέσει πολυδάπανες και απρόβλεπτες χρονικές καθυστερήσεις.
- **Οικονομία:** Οικονομική ανάλυση, πρόβλεψη τιμών συναλλάγματος.
- **Γεωλογικές έρευνες:** Εντοπισμός πετρελαίου και φυσικού αερίου
- **Ρομποτική:** Έλεγχος τροχιάς και σύστημα όρασης ρομπότ
- **Τηλεπικοινωνίες:** Συμπίεση εικόνας και δεδομένων, αυτοματοποιημένες υπηρεσίες πληροφοριών, μετάφραση πραγματικού χρόνου, συστήματα επεξεργασίας πληρωμών
- **Επεξεργασία φωνής:** Αναγνώριση φωνής, συμπίεση φωνής, σύνθεση φωνής από κείμενο
- **Μεταφορές:** Συστήματα διάγνωσης βλαβών φρένων, χρονοπρογραμματισμός οχημάτων, συστήματα δρομολόγησης
- **Ιατρική:** Ένα ευρύ φάσμα ιατρικά συσχετιζόμενων ενδείξεων, όπως ο συνδυασμός της καρδιακής συχνότητας, τα επίπεδα των διαφόρων ουσιών στο αίμα, ο ρυθμός της αναπνοής μπορούν να παρακολουθηθούν. Η εκδήλωση μιας συγκεκριμένης ιατρικής κατάστασης, γίνεται να συσχετιστεί με ένα πολύπλοκο συνδυασμό μεταβολών σε ένα υποσύνολο μεταβλητών που παρακολουθούνται. Τα νευρωνικά δίκτυα έχουν χρησιμοποιηθεί για την αναγνώριση αυτού του προτύπου πρόβλεψης, ώστε να χορηγηθεί η κατάλληλη θεραπεία.
- **Χρηματιστηριακές προβλέψεις:** Οι διακυμάνσεις των τιμών των μετοχών και των χρηματιστηριακών δεικτών είναι ακόμα ένα παράδειγμα ενός πολύπλοκου, πολυδιάστατου φαινομένου. Τα νευρωνικά δίκτυα χρησιμοποιούνται από πολλούς τεχνικούς αναλυτές, ώστε να κάνουν προβλέψεις σχετικά με τις τιμές των μετοχών, βασιζόμενοι σε ένα μεγάλο αριθμό παραγόντων, όπως δηλαδή, τις προηγούμενες επιδόσεις άλλων αποθεμάτων και διαφόρων οικονομικών δεικτών.

ΚΕΦΑΛΑΙΟ 4: Το μοντέλο EMD-BPN

4.1. Εισαγωγή

Σε αυτή τη μελέτη εφαρμόζεται το μοντέλο EMD-BPN. Είναι ένα μοντέλο πρόβλεψης που συνδυάζει τα EMD και BPN. Η σχηματική παράσταση του μοντέλου παρουσιάζεται στο παρακάτω σχήμα.



Σχήμα 17: Μοντέλο EMD-BPN

Σύμφωνα με αυτή τη μέθοδο, πρώτα εισάγουμε τα αρχικά δεδομένα στον αλγόριθμο της EMD για να τα αποσυνθέσουμε σε ένα πεπερασμένο σύνολο εσωτερικών λειτουργιών (IMFs) με απλούστερη συχνότητα, και σε κάποιο υπόλοιπο. Έπειτα, κάθε IMF μοντελοποιείται χρησιμοποιώντας το BPN το οποίο παράγει και την τιμή πρόβλεψης. Οι κρυμμένες πληροφορίες που περιέχουν τα IMFs για τις μεταβλητές πρόβλεψης χρησιμοποιούνται ως μεταβλητές εισόδου στο BPN. Τέλος, η τελική προβλεπόμενη τιμή υπολογίζεται από το άθροισμα των παραπάνω προβλεπόμενων τιμών. Έτσι, μέσω του EMD, διαφορετικά χαρακτηριστικά πληροφορίας, των πρώτων δεδομένων, μπορεί να εμφανίζονται σε διαφορετικές κλίμακες, και η προτεινόμενη μέθοδος μπορεί να συλλάβει πληρέστερα τις τοπικές διακυμάνσεις των πρώτων δεδομένων. Επιπλέον, κάθε IMF έχει παρόμοια συχνότητα χαρακτηριστικών, απλά συστατικά συχνότητας και ισχυρή ομαλότητα. Ως εκ τούτου, μειώνεται η πολυπλοκότητα του BPN, πράγμα που βελτιώνει την ακρίβεια της πρόβλεψης.

4.2. Εφαρμογή της μεθόδου EMD-BPN στην πρόβλεψη ζήτησης τουρισμού

Η διακύμανση και η πολυπλοκότητα της τουριστικής βιομηχανίας καθιστά πολύ δύσκολη την πρόβλεψη της τουριστικής ζήτησης. Σε αυτή τη μελέτη γίνεται μια προσπάθεια πρόβλεψης εφαρμόζοντας τη μέθοδο EMD-BPN. Δηλαδή, ένα μοντέλο πρόβλεψης που πρώτα χρησιμοποιεί την EMD για να αποσυνθέσει τα περίπλοκα αρχικά δεδομένα σε ένα πεπερασμένο σύνολο εσωτερικών λειτουργιών (IMFs) με απλούστερη συχνότητα, και σε κάποιο υπόλοιπο, και έπειτα χρησιμοποιεί τα IMFs και το υπόλοιπο, ως εισόδους για τα νευρωνικά δίκτυα στα οποία γίνεται η μοντελοποίηση και η πρόβλεψη των αποτελεσμάτων. Η τελική τιμή πρόβλεψης είναι το άθροισμα όλων αυτών των αποτελεσμάτων.

Εισαγωγή

Σε ολόκληρο τον κόσμο, η τουριστική βιομηχανία παρουσιάζει πολύ μεγάλη ανάπτυξη τα τελευταία χρόνια. Αυτό έχει ως συνέπεια, την μεγάλη επίδραση των τουριστικών δαπανών στην οικονομική δραστηριότητα κάθε χώρας. Απαιτείται λοιπόν οι προβλέψεις που αφορούν την τουριστική ζήτηση να είναι πιο ακριβείς ώστε να μειωθεί η αβεβαιότητα και το ρίσκο.

Για την πρόβλεψη της τουριστικής ζήτησης έχει παρατηρηθεί σε αρκετές μελέτες η χρήση τεχνητών νευρωνικών δικτύων (ANN), λόγω της ικανότητάς τους να συλλαμβάνουν λεπτές λειτουργικές σχέσεις στο εσωτερικό των εμπειρικών δεδομένων.

Αντίθετα με τα παραδοσιακά στατιστικά μοντέλα, τα νευρωνικά δίκτυα είναι καθοδηγούμενα από δεδομένα και μη παραμετρικά μοντέλα. Μπορούν δηλαδή να χαρτογραφήσουν κάθε γραμμική λειτουργία χωρίς να απαιτούν προηγούμενη αξίωση για τις ιδιότητες των στοιχείων.

Επιπλέον, τα εμπειρικά αποτελέσματα υποδεικνύουν ότι τα νευρωνικά δίκτυα (BPN) με ανάδραση, παρουσιάζουν μεγαλύτερη ακρίβεια στα αποτελέσματα πρόβλεψης, σε σχέση με τα μοντέλα παλινδρόμησης, τα μοντέλα χρονοσειρών και τα εμπρός-τροφοδοτούμενα νευρωνικά δίκτυα.

Η τουριστική βιομηχανία επηρεάζεται και από παράγοντες που δεν μπορούν να προβλεφθούν, όπως για παράδειγμα πολιτικούς παράγοντες και οικονομικά γεγονότα. Δεν είναι εύκολο λοιπόν να περιγράψουμε με ακρίβεια την κυμαινόμενη τάση.

Το empirical mode decomposition προτάθηκε πρώτα από Huang et al.(1998) και είναι το πλέον κατάλληλο για την ανάλυση μη γραμμικού και μη στατικού σήματος, που αναπαριστά με μεγάλη προσαρμοστικότητα τα τοπικά χαρακτηριστικά του δοθέντος σήματος.

Χρησιμοποιώντας την EMD μπορούμε να απλοποιήσουμε οποιοδήποτε πολύπλοκο σήμα σε ένα πεπερασμένο και συνήθως μικρό αριθμό εγγενών λειτουργικών συναρτήσεων (IMFs) . Αυτά περιέχουν τμήματα με απλούστερη συχνότητα και μεγαλύτερη αλληλεξάρτηση, πράγμα που διευκολύνει την πρόβλεψή τους. Για αυτό το λόγο εφαρμόζεται ευρέως σε πολλά πεδία, όπως για παράδειγμα στην ανάλυση μη στατικών θαλάσσιων κυμάτων, σε σεισμικά σήματα και δομικές αναλύσεις κ.α..

Σε αυτή την εργασία χρησιμοποιείται η μέθοδος EMD-BPN για να προβλεφθεί ο αριθμός αφίξεων στην Ταϊβάν. Για να αξιολογηθεί η αποτελεσματικότητα της συγκεκριμένης προσέγγισης, συγκρίθηκαν τα αποτελέσματα, με αυτά των απλών νευρωνικών δικτύων, και των παραδοσιακών χρονικών σειριακών μοντέλων όπως το autoregressive integrated moving average(ARIMA). Από αυτήν τη σύγκριση αποδείχθηκε ότι το προτεινόμενο μοντέλο παρουσιάζει άψογη ικανότητα πρόβλεψης.

Το προτεινόμενο μοντέλο EMD-BPN

Σε αυτή τη μελέτη εφαρμόζεται το μοντέλο EMD-BPN.

Πρώτα, εισάγονται τα αρχικά δεδομένα της τουριστικής ζήτησης στον EMD για να αποσυντεθούν σε ένα πεπερασμένο αριθμό IMFs. (Εδώ το υπόλειμμα r_n αντιμετωπίζεται ως IMF).

Δεύτερον, κάθε IMF μοντελοποιείται χρησιμοποιώντας το BPN το οποίο παράγει και την τιμή πρόβλεψης.

Τέλος, η τελική προβλεπόμενη τιμή υπολογίζεται από το άθροισμα των παραπάνω προβλεπόμενων τιμών.

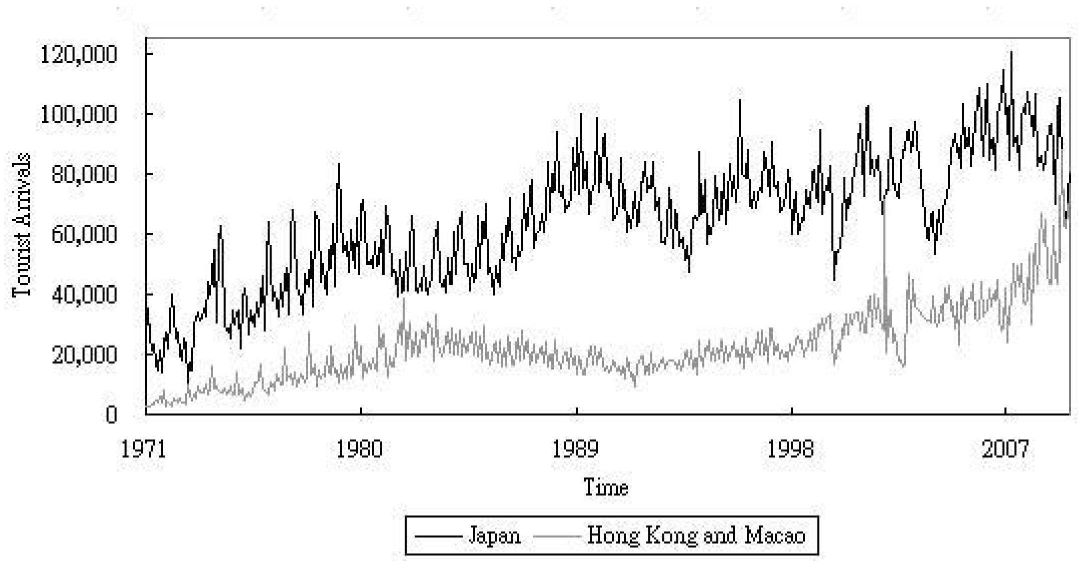
Οι κρυμμένες πληροφορίες που περιέχουν τα IMFs για τις μεταβλητές πρόβλεψης χρησιμοποιούνται ως μεταβλητές εισόδου στο BPN. Σχεδιάστηκε ένα μοντέλο BPN με ένα μόνο κρυμμένο στρώμα, εφόσον ένα κρυμμένο στρώμα αρκεί, για να μοντελοποιήσουμε κάθε περίπλοκο σύστημα με μεγάλη ακρίβεια. Δεν υπάρχουν γενικοί κανόνες για αν υπολογίσουμε τον βέλτιστο αριθμό κόμβων που απαιτείται. Έτσι, σε αυτή την μελέτη καθορίστηκε η βέλτιστη τοπολογία σύμφωνα με τη μέθοδο «δοκιμή-σφάλμα».

Επιπλέον, κάθε IMF έχει παρόμοια συχνότητα χαρακτηριστικών, απλά συστατικά συχνότητας και ισχυρή ομαλότητα. Ως εκ τούτου, μειώνεται η πολυπλοκότητα του BPN, πράγμα που βελτιώνει την ακρίβεια της πρόβλεψης.

Πειραματικά αποτελέσματα

Σύνολα δεδομένων

Σε αυτή τη μελέτη χρησιμοποιήθηκαν ως δεδομένα, οι μηνιαίες αφίξεις στην Ταϊβάν, από τον Ιανουάριο του 1971, έως τον Αύγουστο του 2007. Συνολικά είναι 459 σημεία δεδομένων. Η μηνιαία σειρά απεικονίζεται στο σχήμα 18.



Σχήμα 18

Παρατηρούμε ότι οι σειρές έχουν ανοδική πορεία με βραχυπρόθεσμες διακυμάνσεις.

Για να αξιολογηθούν οι επιδόσεις του προτεινόμενου μοντέλου πρόβλεψης, χωρίστηκαν τα συλλεγόμενα δεδομένα σε δύο ομάδες, το δείγμα εκπαίδευσης και το δείγμα δοκιμών. Πιο συγκεκριμένα, τα πρώτα 368 σημεία των δεδομένων (79,98% των συνολικών δεδομένων) αποτελούν το δείγμα εκπαίδευσης. Αντίστοιχα τα υπόλοιπα 91 σημεία δεδομένων (20% των συνολικών δεδομένων) αποτελούν το δείγμα ελέγχου. Δηλαδή, το διάστημα από τον Ιανουάριο του 1971 έως τον Αύγουστο του 2001 έχει επιλεγεί ως περίοδος κατάρτισης και από τον Σεπτέμβριο του 2001 έως τον Αύγουστο του 2009 είναι η περίοδος δοκιμών. Η μέθοδος *ένα-βήμα-μπροστά* παρουσιάζεται στα πειράματα.

Απόδοση κριτηρίων

Για την αξιολόγηση των αποτελεσμάτων χρησιμοποιήθηκαν κάποια κριτήρια. Το κριτήριο του μέσου ποσοστιαίου σφάλματος (MAPE), το κριτήριο της μέσης τετραγωνικής απόκλισης (RMSE) και το κριτήριο της απόλυτης διαφορά (MAD). Οι τύποι παρουσιάζονται στον πίνακα 1. Όσο πιο μικρές είναι οι τιμές των σφαλμάτων, τόσο πιο κοντά στις πραγματικές τιμές είναι οι προβλεπόμενες.

Metrics	Calculation
MAPE*	$\frac{1}{n} \times \sum_{i=1}^{n} \left| \frac{P_i - A_i}{P_i} \right| \times 100\%$
RMSE	$\sqrt{\frac{1}{n} \times \sum_{i=1}^{n} (P_i - A_i)^2}$
MAD	$\frac{1}{n} \times \sum_{i=1}^{n} \lvert P_i - A_i \rvert$

*A_i and P_i present the actual and predicted values.

Πίνακας 1

Αποτελέσματα πρόβλεψης

Για να αξιολογηθούν την αποτελεσματικότητα του προτεινόμενου μοντέλου EMD-BPN , συγκρίθηκαν τα αποτελέσματα με αυτά του απλού μοντέλου πρόβλεψης BPN. Το απλό μοντέλο BPN χρησιμοποιεί τον αριθμό τουριστών του προηγούμενου μήνα για να προβλέψει το πλήθος των τουριστών του επόμενου μήνα. Σύμφωνα με τον EMD ο αριθμός των τουριστών από την Ιαπωνία αποσυντίθεται σε 7 IMFs κι ένα υπόλειμμα, το οποίο παρουσιάζει μια διαρκή και ομαλή μεταβολή. Αυτό σημαίνει ότι η διακοπή και η σύνδεση μεταξύ των διαφορετικών χαρακτηριστικών των πληροφοριών που βρίσκονται στα αρχικά δεδομένα, έχουν αποδυναμωθεί σε κάποιο βαθμό. Έτσι, είναι πιο εύκολο να χτιστεί το μοντέλο πρόβλεψης.

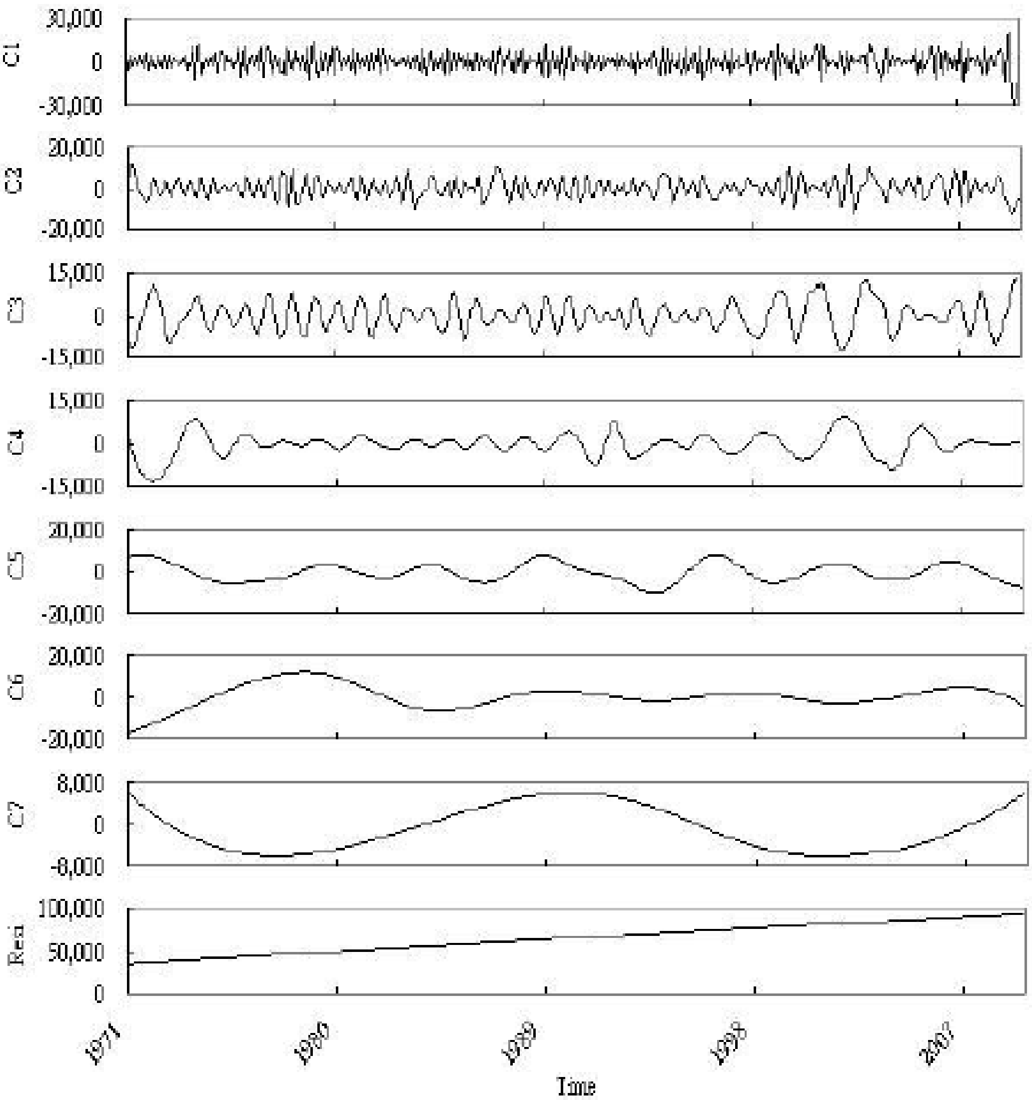


Figure 3. Empirical mode decomposition for visitor arrivals from Japan

<u>Σχήμα 19</u>

Εφόσον μετά την χρήση του EMD καταλήξαμε σε 7 IMFs και σε ένα υπόλειμμα, το μοντέλο BPN θα έχει 8 κόμβους εισόδου. Δηλαδή 8 μεταβλητές πρόβλεψης. Για να βρεθεί ο αριθμός των κρυμμένων κόμβων έγιναν κάποιες δοκιμές.

Ο αριθμός των κόμβων που θα συγκριθούν είναι 14,15,16,17 και 18. Το δίκτυο έχει μόνο έναν κόμβο εξόδου, τον αριθμό αφίξεων. Μιας και οι μικρότεροι ρυθμοί μάθησης τείνουν να δίνουν καλύτερα αποτελέσματα, δοκιμάζουμε τους ρυθμούς μάθησης 0,02 , 0,025 και 0,03 κατά την διαδικασία κατάρτισης.

Στην μοντελοποίηση του απλού μοντέλου BPN τα στρώματα εισόδου και εξόδου έχουν από έναν κόμβο. Ο αριθμός των κόμβων στο κρυφό στρώμα προς δοκιμή είναι 14,15,16 και 17. Οι ρυθμοί μάθησης που δοκιμάζονται κατά την διάρκεια κατάρτισης είναι 0,02 , 0,025 και 0,03.

Ως κριτήριο σύγκλισης για τα δύο μοντέλα, χρησιμοποιείται το κριτήριο του μέσου τετραγωνικού σφάλματος (RMSE), με ρίζα μικρότερη ή ίση του 0,0001 ή το μέγιστο όριο των 10000 επαναλήψεις. Η τοπολογία δικτύου με τις ελάχιστες δοκιμές RMSE θεωρείται ως βέλτιστο δίκτυο. Τα αποτελέσματα των δύο μοντέλων, ελέγχοντας όλες τους δυνατούς συνδυασμούς κρυφών κόμβων και ρυθμών μάθησης, παρουσιάζονται στους πίνακες 2 και 3.

Model selection results of the EMD-BPN model for visitor arrivals from Japan

Number of nodes in the hidden layer	Learning rate	Training RMSE	Testing RMSE
14	0.02	0.00281	0.00953
	0.025	0.00266	0.00989
	0.03	0.00256	0.01010
15	**0.02**	**0.00294**	**0.00773**
	0.025	0.00277	0.00828
	0.03	0.00264	0.00881
16	0.02	0.00302	0.00921
	0.025	0.00285	0.00988
	0.03	0.00272	0.01043
17	0.02	0.00268	0.00909
	0.025	0.00259	0.00953
	0.03	0.00252	0.00989
18	0.02	0.00282	0.00979
	0.025	0.00268	0.01027
	0.03	0.00258	0.01067

Πίνακας 2

Table 3

Model selection results of the single BPN model for visitor arrivals from Japan

Number of nodes in the hidden layer	Learning rate	Training RMSE	Testing RMSE
1	0.02	0.00534	0.01671
	0.025	0.00514	0.01609
	0.03	0.00499	0.01558
2	0.02	0.00436	0.01353
	0.025	0.00421	0.01305
	0.03	0.00408	0.01264
3	0.02	0.00402	0.01259
	0.025	0.00391	0.01221
	0.03	**0.00381**	**0.01189**
4	0.02	0.00427	0.01309
	0.025	0.00410	0.01260
	0.03	0.00396	0.01218

Πίνακας 3

Σύμφωνα με τον πίνακα 2, για το μοντέλο EMD-BPN η βέλτιστη τοπολογία είναι {8,15,1} με ρυθμό μάθησης 0,02, εφόσον αυτή δίνει καλύτερο αποτέλεσμα πρόβλεψης (ελάχιστη δοκιμασία RMSE). Δηλαδή, 8 κόμβους εισόδου, 15 κρυφούς κόμβους και 1 κόμβο εξόδου.

Σύμφωνα με τον πίνακα 3, για το απλό μοντέλο BPN, η βέλτιστη τοπολογία είναι {1,16,1} με ρυθμό μάθησης 0,03.

Τα μοντέλα ARIMA συμπεριλαμβάνονται ως μοντέλα αναφοράς στη σύγκριση των επιδόσεων πρόβλεψης. Το τυπικό εποχικό-μη-εποχικό μοντέλο ARIMA, συμβολίζεται με ARIMA(p,d,q)(P,D,Q), όπου το p και το P αντιπροσωπεύουν τις εντολές των εποχικών και μη εποχικών αυτοπαλίνδρομων παραμέτρων αντίστοιχα, τα d και D αντιπροσωπεύουν τον αριθμό των τακτικών και εποχικών διαφορών αντίστοιχα που απαιτούνται, και τα q και Q αντιπροσωπεύουν τις εντολές των μη εποχικών και εποχικών, με κινητό μέσο όρο, παραμέτρων αντίστοιχα.

Σε αυτή τη μελέτη το μοντέλο ARIMA έχει τρεις φάσεις που βοηθούν στον υπολογισμό του καταλληλότερου μοντέλου. Προσδιορισμό μοντέλου, εκτίμηση παραμέτρων και διαγνωστικό έλεγχο. Τελικά ως καλύτερο μοντέλο προκύπτει το ARIMA(2,1,1)(1,1,1). Στον πίνακα 4 φαίνονται οι τιμές που παίρνουν τα σφάλματα MAPE, RMSE και MAD για τα τρία μοντέλα πρόβλεψης. Παρατηρούμε ότι το μοντέλο EMD-BPN παρουσιάζει τις χαμηλότερες τιμές σφάλματος (MAPE:0.958%, RMSE:1443.1 MAD:900,8), πράγμα που σημαίνει ότι παρέχει καλύτερα αποτελέσματα πρόβλεψης από το απλό BPN και από το ARIMA.

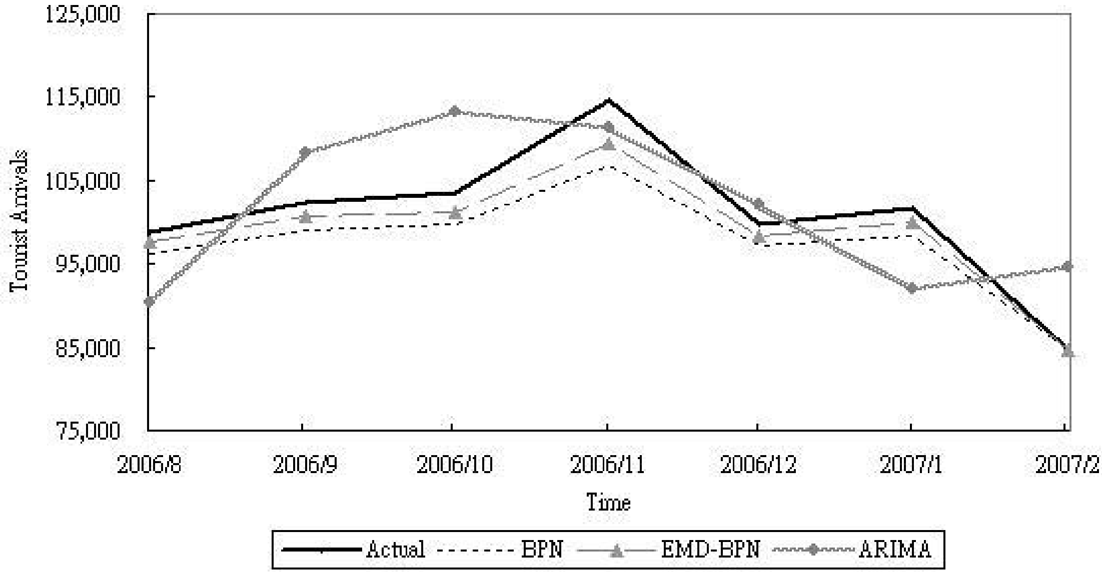
Table 4

Performance of the three forecasting methods for visitor arrivals from Japan

Metrics	Model		
	EMD-BPN	Single BPN	ARIMA(2,1,1)(1,1,1)12
MAPE	0.958 %	1.378 %	8.876 %
RMSE	1443.1	2237.26	9859.8
MAD	900.08	1360.17	7569.87

Πίνακας 4

Στον σχήμα 20 παρουσιάζονται οι πραγματικές αφίξεις επισκεπτών από την Ιαπωνία, καθώς και οι προβλεπόμενες με τα τρία διαφορετικά μοντέλα πρόβλεψης. Μπορούμε να παρατηρήσουμε ότι οι τιμές πρόβλεψης του μοντέλου EMD-BPN είναι πιο κοντά στις πραγματικές, σε σχέση με αυτές των άλλων μοντέλων.



Σχήμα 20

Εφαρμόζοντας την ίδια διαδικασία για τις αφίξεις από το Χονγκ Κονγκ και το Μακάο, παρατηρούμε ότι και πάλι το μοντέλο πρόβλεψης EMD-BPN παρουσιάζει τις μικρότερες τιμές σφάλματος.

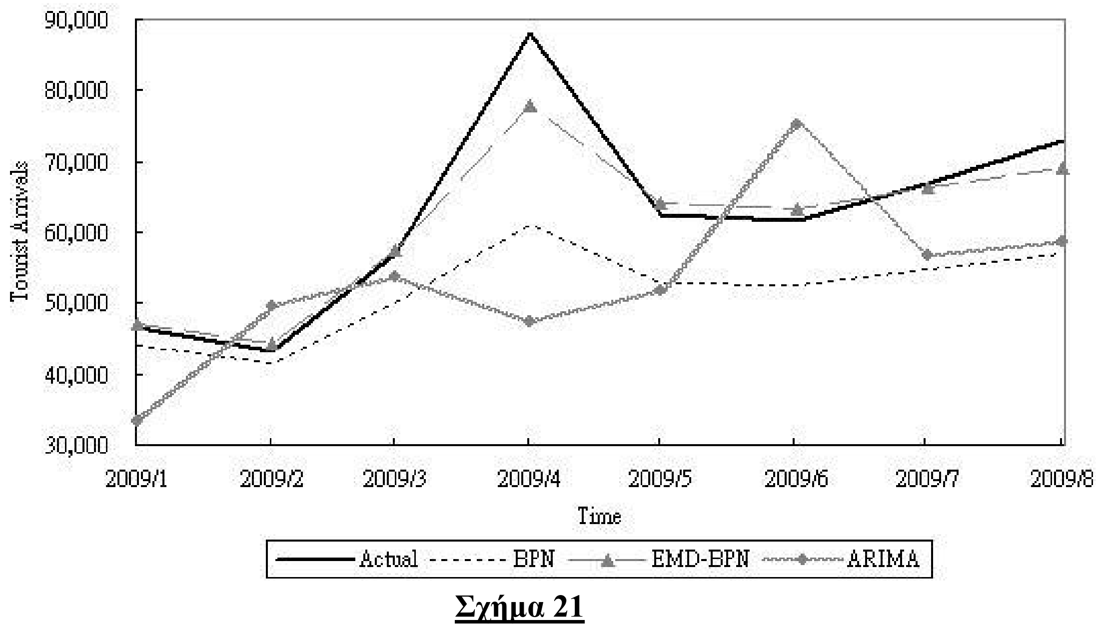
Table 5

Performance of three forecasting methods for visitor arrivals from Hong Kong and Macao

Metrics	Model		
	EMD-BPN	Single BPN	ARIMA(2,1,0)(1,1,1)12
MAPE	2.741%	4.049%	18.193 %
RMSE	1707.11	4917	9929.36
MAD	1192.94	2324.51	7082.25

Πίνακας 5

Στον σχήμα 21 παρουσιάζονται οι πραγματικές αφίξεις επισκεπτών από το Χονγκ Κονγκ και το Μακάο, καθώς και οι προβλεπόμενες με τα τρία διαφορετικά μοντέλα πρόβλεψης. Μπορούμε να παρατηρήσουμε ότι οι τιμές πρόβλεψης του μοντέλου EMD-BPN είναι και πάλι πιο κοντά στις πραγματικές, σε σχέση με αυτές των άλλων μοντέλων.



Σχήμα 21

Συμπεράσματα

Το έγγραφο παρουσίασε ένα μοντέλο πρόβλεψης της τουριστικής ζήτησης, που βασίζεται στα EMD και BPN. Το μοντέλο EMD-BPN είναι κατάλληλο για την πρόβλεψη μη στάσιμων δεδομένων. Η διαδικασία που ακολουθεί αυτή η μέθοδος είναι η εξής:

Πρώτα αποσυνθέτει να αρχικά δεδομένα σε έναν πεπερασμένο αριθμό IMFs και σ' ένα υπόλειμμα, δηλαδή σε τμήματα με απλούστερη συχνότητα και δυνατότερη αλληλεξάρτηση, που είναι πιο εύκολο να μοντελοποιηθούν σε σχέση με τα αρχικά δεδομένα.

Έπειτα, κάθε ένα από αυτά τα τμήματα γίνεται είσοδος για τον BPN, που μοντελοποιεί αυτά τα δεδομένα και καταλήγει σε μία τιμή πρόβλεψης.

Στο τέλος αθροίζονται όλες οι τιμές πρόβλεψης και μας δίνουν την τελική τιμή πρόβλεψης.

Τα πειράματα της εργασίας αξιολόγησαν τις αφίξεις στην Ταϊβάν από την Ιαπωνία, και από το Χονγκ Κονγκ και Μακάο. Χρησιμοποίησε τα κριτήρια MAPE, RMSE, και MAD για να συγκρίνει την μέθοδο EMD-BPN με τις μεθόδους BPN και ARIMA. Σύμφωνα με τα αποτελέσματα, η μέθοδος EMD-BPN παρουσιάζει μικρότερες τιμές σφάλματος από τις άλλες μεθόδους και οι προβλεπόμενες τιμές που δίνει οι πιο κοντά στις πραγματικές, απ' ότι οι τιμές των άλλων μεθόδων.

Επομένως, η μέθοδος EMD-BPN είναι η πιο κατάλληλη για την επεξεργασία μη γραμμικών και μη στατικών δεδομένων με ισχυρή πολυπλοκότητα και είναι μια αποτελεσματική μέθοδος για την πρόβλεψη της τουριστικής ζήτησης.

ΚΕΦΑΛΑΙΟ 5: Μελέτη περίπτωσης

Στόχος της παρούσας εργασίας είναι να ελέγξουμε τα αποτελέσματα της μοντέλου EMD-BPN στην πρόβλεψη των τιμών κάποιων προϊόντων.

Στο πρώτο βήμα της εργασίας είναι η συλλογή των δεδομένων. Πιο συγκεκριμένα, συγκεντρώθηκαν οι τιμές κάποιων προϊόντων για όλους τους μήνες από τον Ιανουάριο του 1969 έως τον Οκτώβριο του 2010, δηλαδή 502 τιμές για κάθε προϊόν. Τα προϊόντα που χρησιμοποιήθηκαν είναι τα εξής:

- Σιτάρι (Wheat)
- Ζάχαρη (Sugar)
- Καφές (Coffee)
- Κακάο (Cocoa)
- Σόγια (Soybeans)

Στον παρακάτω πίνακα παρουσιάζονται αναλυτικά τα δεδομένα:

PRODUCT	**Wheat, United States, n° 2 Hard Red Winter (ordinary), FOB Gulf**	**Sugar, average of I.S.A. daily prices, FOB Caribbean ports (Ά/lb.)**	**Coffee, Robustas, ex-dock New York (Ά/lb.)**	**Cocoa beans, average daily prices New York/London (Ά/lb.)**	**Soybeans, in bulk, United States, n° 2 yellow, CIF Rotterdam**
PERIOD					
Jan1969	63,2	2,88	32	42,54	111

Feb1969	62,8	3,12	32,15	41,3	110
Mar1969	62,1	3,51	31,24	40,75	110
Apr1969	62,1	3,65	29,92	39,73	110
May1969	62,5	3,65	29,47	39,69	113
Jun1969	61,7	3,64	30,7	40,5	110
Jul1969	59,9	3,47	31,17	42,24	101
Aug1969	56,2	2,95	32,87	41,47	99
Sep1969	53,6	2,74	34,61	40,86	101
Oct1969	52,5	2,95	38,33	41,68	103
Nov1969	52,5	3,04	37,23	42,28	107
Dec1969	52,2	2,82	37,58	38,61	106
Jan1970	52,2	3,06	39,11	33,86	110
Feb1970	52,2	3,15	38,45	30,67	113
Mar1970	51,8	3,38	39,37	30,38	113
Apr1970	51,8	3,57	41,41	29,8	113
May1970	52,2	3,61	42,64	27,3	117
Jun1970	52,2	3,69	42,18	26,68	121
Jul1970	52,9	3,82	42,75	28,91	127
Aug1970	54	3,81	42,85	33,46	122
Sep1970	56,2	3,87	42,58	33,85	123
Oct1970	58,8	3,93	42,71	32,06	130
Nov1970	60,6	4,09	41,55	30,38	132
Dec1970	63,2	4,11	41,66	29,48	128
Jan1971	63,6	4,72	42,18	27,23	129
Feb1971	63,6	4,82	42,27	25,48	129
Mar1971	62,8	4,69	42,67	24,47	128
Apr1971	62,5	4,56	42,64	24,76	124
May1971	63,2	4,36	42,64	23,16	127
Jun1971	61,4	4,15	42,37	24,04	132
Jul1971	61,7	4,17	42,11	25,89	136
Aug1971	61	4,25	42,18	26,91	131
Sep1971	61	3,99	41,93	25	131
Oct1971	60,3	4,21	41,69	22,47	133
Nov1971	59,9	4,24	41,98	21,96	132
Dec1971	59,9	5,78	42,52	21,79	134
Jan1972	59,9	7,9	42,17	24,26	132
Feb1972	59,9	8,19	42,22	24,94	134
Mar1972	59,9	8,4	43,07	26,12	143
Apr1972	60,3	7,08	43,67	26,37	146
May1972	60,3	6,63	43,76	27,87	146
Jun1972	59,9	6,33	43,66	28,19	142
Jul1972	59,9	5,56	47,35	29,24	137
Aug1972	62,1	6,26	47,2	31,04	138
Sep1972	76,1	7,07	46,93	32,85	144
Oct1972	87,4	7,41	47,03	33,02	146
Nov1972	88,9	7,28	47,4	32,51	150
Dec1972	102,9	9,15	47,77	33,37	168

Jan1973	108,4	9,4	47,82	32,47	214
Feb1973	98,8	8,98	49,53	33,87	259
Mar1973	98,5	8,77	50,34	38,3	258
Apr1973	100,7	8,99	48,85	42,73	260
May1973	102,1	9,35	48,76	53,03	358
Jun1973	108,4	9,38	48,91	58,02	470
Jul1973	114,6	9,52	47,99	70,3	330
Aug1973	174,2	8,97	47,36	64,22	331
Sep1973	191,4	8,94	50,08	60,25	266
Oct1973	181,5	9,51	52	58,72	245
Nov1973	179,7	10,14	52,97	51,87	239
Dec1973	196,9	11,85	53,95	51,74	254
Jan1974	216	15,16	55,65	50,89	261
Feb1974	221	21,09	60,55	58,38	271
Mar1974	199	21,1	62,72	67,39	265
Apr1974	159	21,6	64,03	82,6	235
May1974	142	23,63	64,53	82,71	229
Jun1974	154	23,51	61,8	70,28	228
Jul1974	170	25,03	59,13	71,67	276
Aug1974	170	30,63	56,01	73,48	320
Sep1974	172	34,15	53,92	72,35	303
Oct1974	193	39,5	54,94	79,18	335
Nov1974	191	56,14	55,66	75,59	312
Dec1974	186	44,88	55,18	64,36	288
Jan1975	169	38,31	54,32	67,19	256
Feb1975	158	33,98	52,39	67,84	231
Mar1975	147	26,4	49,64	62,64	226
Apr1975	147	23,95	49,03	56,2	229
May1975	133	17,37	47,52	48,01	208
Jun1975	127	13,65	49,34	44,9	207
Jul1975	143	16,69	58,68	52,85	224
Aug1975	160	18,61	77,05	53,25	243
Sep1975	168	15,5	76,1	52,47	227
Oct1975	166	14,07	73,07	56,7	211
Nov1975	152	13,47	71,41	55,99	192
Dec1975	144	13,19	74,03	59,83	185
Jan1976	143	14,02	79,13	63,7	189
Feb1976	154	13,5	82,69	65,04	191
Mar1976	155	14,79	82,41	64,17	190
Apr1976	147	14,05	107,79	64,32	190
May1976	143	14,54	121,64	80,2	210
Jun1976	147	12,99	130,06	87,26	248
Jul1976	143	13,21	126,29	88,75	264
Aug1976	128	10,02	132,26	96,23	248
Sep1976	124	8,13	137,08	108,41	261
Oct1976	113	8,03	151,9	117,59	256
Nov1976	109	7,88	176,26	132,67	260

Dec1976	108	7,55	203,96	134,92	269
Jan1977	109	8,34	216,26	156,45	287
Feb1977	112	8,59	246,37	173,15	293
Mar1977	110	8,98	306,09	182,92	328
Apr1977	104	10,04	312,24	162,66	384
May1977	98	8,95	269,56	171,53	371
Jun1977	93	7,87	224,02	192,84	326
Jul1977	98	7,39	196,49	197,84	252
Aug1977	96	7,61	203,33	180,02	230
Sep1977	102	7,31	201,5	174,65	205
Oct1977	106	7,09	173,93	167,48	209
Nov1977	114	7,07	166,71	157,87	236
Dec1977	116	8,09	168,57	144,87	241
Jan1978	119	8,77	177,13	135,08	239
Feb1978	120	8,48	174,56	128,14	239
Mar1978	127	7,74	157,64	153,35	273
Apr1978	135	7,59	144,8	156,1	290
May1978	128	7,33	136,09	142,19	290
Jun1978	130	7,23	151,14	135,68	278
Jul1978	130	6,43	126,54	143,47	266
Aug1978	129	7,08	128,45	152,43	262
Sep1978	134	8,17	146,9	170,09	264
Oct1978	139	8,96	150,12	173,57	271
Nov1978	141	8,01	144,26	182,77	270
Dec1978	139	8	132,08	179,03	278
Jan1979	140	7,57	133,25	165,77	284
Feb1979	144	8,23	132,5	158,85	298
Mar1979	143	8,46	135,2	154,02	310
Apr1979	141	7,82	141,8	147,29	300
May1979	146	7,85	147,68	152,32	300
Jun1979	164	8,14	188,79	157,82	322
Jul1979	179	8,52	195,9	147,47	322
Aug1979	173	8,85	181,94	142,18	302
Sep1979	179	9,9	189,51	146,13	292
Oct1979	183	11,94	184,34	138,72	283
Nov1979	183	13,68	177,75	136,77	281
Dec1979	185	14,93	176,97	144,52	279
Jan1980	179	17,16	162,56	143,63	268
Feb1980	176	22,75	162	146,78	271
Mar1980	169	19,64	169,89	140,18	264
Apr1980	158	21,25	162,9	131,97	252
May1980	163	30,94	174,06	117,24	260
Jun1980	160	30,8	169,01	113,27	262
Jul1980	169	27,7	147,94	111,77	303
Aug1980	175	31,77	132,05	104,85	309
Sep1980	183	34,74	124,96	108,2	335
Oct1980	192	40,55	125,49	104,54	340

Nov1980	200	37,81	116,36	98,57	367
Dec1980	188	28,79	118,54	95,7	324
Jan1981	193	27,78	121,88	94,94	323
Feb1981	185	24,09	115,23	92,57	306
Mar1981	176	21,81	113,92	95,97	305
Apr1981	181	17,83	112,93	95,69	316
May1981	174	15,06	108,38	87,68	306
Jun1981	170	16,38	85,23	75,31	291
Jul1981	171	16,34	88,25	92,91	294
Aug1981	173	14,76	90,39	101,92	283
Sep1981	174	11,65	88,88	106,01	264
Oct1981	173	12,04	98,51	98,55	260
Nov1981	180	11,97	106,33	93,28	257
Dec1981	182	12,98	104,52	95,46	256
Jan1982	175	12,9	106,01	99,25	262
Feb1982	172	13,08	115,63	94,67	254
Mar1982	169	11,26	112,83	88,35	254
Apr1982	170	9,58	106,67	79,17	265
May1982	160	8,11	102,96	77,79	269
Jun1982	152	6,84	102,85	71,29	254
Jul1982	155	7,8	99,67	71,59	249
Aug1982	155	6,77	103,79	70,74	234
Sep1982	158	5,77	112	76,48	216
Oct1982	151	5,93	118,27	74,3	214
Nov1982	157	6,52	123,15	69,91	231
Dec1982	164	6,31	128,76	74,55	232
Jan1983	166	6,03	124,71	80,93	235
Feb1983	166	6,43	122,15	88,1	239
Mar1983	167	6,2	119,99	84,34	242
Apr1983	168	6,71	120,42	85,15	253
May1983	157	9,24	122,64	93,67	252
Jun1983	151	10,74	120,55	104,22	243
Jul1983	148	10,53	121,01	104,45	263
Aug1983	154	10,56	120,9	104,46	339
Sep1983	159	9,43	122,1	96,98	350
Oct1983	154	9,69	130,34	95	329
Nov1983	153	8,33	130,38	100,68	324
Dec1983	154	7,67	134,21	115,25	311
Jan1984	154	6,97	134,24	118,93	305
Feb1984	151	6,63	137,19	113,28	293
Mar1984	155	6,43	138,88	115,04	314
Apr1984	158	6	137,58	115,05	315
May1984	154	5,61	146,24	121,45	338
Jun1984	150	5,52	143,67	112,72	308
Jul1984	149	4,54	138,32	100,54	270
Aug1984	155	4,05	140,11	100,41	261
Sep1984	156	4,1	142,05	104,83	245

Oct1984	155	4,65	135,01	102,39	245
Nov1984	153	4,36	135,47	102,29	250
Dec1984	150	3,55	129,13	97,09	241
Jan1985	152	3,61	127,9	100,68	243
Feb1985	149	3,7	124,67	103,12	239
Mar1985	146	3,83	124,02	101,26	241
Apr1985	146	3,42	123,25	101,64	243
May1985	138	2,88	121,41	96,57	231
Jun1985	134	2,78	120,52	94,88	227
Jul1985	131	3,18	107,29	100,24	223
Aug1985	125	4,39	106,81	103,21	221
Sep1985	128	5,12	105,11	106,24	207
Oct1985	131	5,01	112,48	107,19	205
Nov1985	135	5,48	128,03	104,22	210
Dec1985	139	5,32	154,9	107,93	213
Jan1986	134	4,84	172,22	105,4	221
Feb1986	131	5,56	163,32	100,63	216
Mar1986	136	7,04	169,79	95,58	217
Apr1986	120	8,33	157,53	90,25	213
May1986	115	7,67	142,56	86,44	215
Jun1986	107	6,34	125,79	86,48	213
Jul1986	103	5,55	126,42	91,74	209
Aug1986	105	5,57	135,17	93,79	202
Sep1986	104	4,68	163,8	100,08	195
Oct1986	108	5,39	148,58	94,03	197
Nov1986	107	5,95	142,71	91,65	204
Dec1986	109	5,71	126,02	89,65	199
Jan1987	112	6,49	113,81	89,66	200
Feb1987	113	7,38	110,45	88,83	198
Mar1987	117	7,56	98,28	91,16	199
Apr1987	115	6,68	101,17	93,88	207
May1987	119	6,73	105,66	93,82	221
Jun1987	110	6,44	98,23	90,74	228
Jul1987	106	6,1	92,5	95,93	223
Aug1987	108	5,62	94,8	91,7	215
Sep1987	114	5,82	99,48	90,16	214
Oct1987	117	6,65	103,83	86,73	217
Nov1987	117	7,33	105,76	87,91	228
Dec1987	126	8,3	104,05	86,05	240
Jan1988	128	9,67	103,19	89,5	254
Feb1988	132	8,43	104,09	81,19	254
Mar1988	125	8,52	99,82	76,44	257
Apr1988	128	8,54	96,6	74,8	270
May1988	133	8,9	94,94	77,42	285
Jun1988	152	10,57	94,15	73,67	353
Jul1988	152	14,02	85,39	72,95	351
Aug1988	151	11,15	81,58	65,67	342

Sep1988	161	10,16	89,98	56,21	337
Oct1988	163	10,28	95,24	59,16	313
Nov1988	164	10,84	94,13	69,96	313
Dec1988	168	11,22	102,17	68,12	313
Jan1989	175	9,65	102,83	65,17	318
Feb1989	172	10,54	96,66	67,23	307
Mar1989	177	11,54	95,61	63,9	313
Apr1989	176	12,16	91,73	58,56	297
May1989	170	11,98	92	54,78	297
Jun1989	169	12,64	84,55	56,8	291
Jul1989	169	13,99	65,78	61,02	272
Aug1989	166	14	59,35	60,16	236
Sep1989	164	14,13	60,66	51,75	237
Oct1989	166	14,44	53,83	48,25	238
Nov1989	168	15,02	53,23	46,67	246
Dec1989	169	13,41	52,15	44,01	248
Jan1990	170	14,2	50,52	45,13	242
Feb1990	162	14,64	51,3	46,56	240
Mar1990	159	15,36	57,68	51,32	242
Apr1990	160	15,24	57,39	62,25	247
May1990	149	14,63	55,03	65,61	259
Jun1990	136	12,97	52,38	60,08	249
Jul1990	126	11,99	51,47	61,04	252
Aug1990	119	10,93	55,1	58,53	250
Sep1990	115,5	11,02	57,09	62,21	250
Oct1990	116,1	9,82	58,05	59,31	247
Nov1990	113,5	10,06	56,79	61,32	239
Dec1990	115,2	9,74	57,04	58,69	244
Jan1991	113,5	8,8	53,92	56,36	239
Feb1991	114,5	8,51	52,46	55,1	241
Mar1991	120,1	9,14	52,13	54,9	244
Apr1991	121	8,53	52,38	51,97	245
May1991	122	7,58	48,22	48,18	241
Jun1991	121	9,2	47,1	47,21	241
Jul1991	117	10,31	46,49	46,75	229
Aug1991	127,1	9,46	46,57	51,64	241
Sep1991	133,7	9,29	48,11	57,8	246
Oct1991	146,3	9,13	47,31	58,97	237
Nov1991	150,6	8,66	51,47	59,68	237
Dec1991	159,6	9	51,45	61,99	234
Jan1992	163,4	8,42	49,44	58,17	236
Feb1992	177,4	7,84	43,22	54,1	237
Mar1992	168,9	8,25	43,08	51,17	240
Apr1992	160,3	9,46	42,32	47,37	235
May1992	149,3	9,57	38,79	45,15	247
Jun1992	148,7	10,36	38,07	43,67	253
Jul1992	137,6	10,31	39,6	50	237

Aug1992	130,1	9,83	39,92	53,43	229
Sep1992	141,6	9,31	42,39	51,03	227
Oct1992	141,1	8,71	45,5	48,58	221
Nov1992	145,1	8,57	49,08	47,8	229
Dec1992	150,1	8,15	52,09	46,98	235
Jan1993	155,6	8,24	48,13	46,97	239
Feb1993	150,1	8,57	48,25	45,22	236
Mar1993	150,8	10,62	46,86	44,24	238
Apr1993	144,2	11,15	45,51	45,92	244
May1993	137,9	11,83	46,91	45,24	246
Jun1993	123,9	10,39	47,65	44,78	245
Jul1993	127	9,68	50,39	48,1	287
Aug1993	133,8	9,34	59,29	49,8	274
Sep1993	134,6	9,53	63,44	58,25	260
Oct1993	138,8	10,29	60,05	58,63	249
Nov1993	148,7	10,12	62,53	60,29	267
Dec1993	155,8	10,52	62,95	61,47	276
Jan1994	158,2	10,3	60,91	57,33	282
Feb1994	146,9	10,82	62,25	55,74	273
Mar1994	142,6	10,3	66,46	59,29	274
Apr1994	142,3	11,01	72,64	56,36	261
May1994	142	11,58	96,05	62,5	256
Jun1994	140,5	12,05	113,31	66,18	265
Jul1994	138,3	11,77	164,65	70,41	237
Aug1994	146	12,1	162,68	70,37	232
Sep1994	158,7	12,59	182,95	67,4	233
Oct1994	168,3	12,75	170,09	65,68	227
Nov1994	163,4	13,94	154,19	65,18	239
Dec1994	164,2	14,67	130,48	63,48	243
Jan1995	159,8	14,76	132,26	66,63	243
Feb1995	155,2	14,41	135,22	68,51	238
Mar1995	149,8	14,59	146,83	67,36	246
Apr1995	151	13,7	145,47	66,64	249
May1995	162,7	13,52	141,89	65,56	250
Jun1995	171,4	14,02	129,53	65,19	252
Jul1995	191,3	13,59	120,89	61,87	261
Aug1995	185,1	12,98	131,28	64,09	257
Sep1995	195,3	11,69	116,41	63,36	266
Oct1995	205,8	11,81	114,15	63,26	271
Nov1995	204,9	11,98	112,79	65,16	282
Dec1995	212,9	12,31	94,72	62,54	296
Jan1996	209,8	12,53	91,99	61,22	305
Feb1996	218,3	12,81	98,99	62,14	299
Mar1996	217,6	12,91	91,99	60,72	295
Apr1996	256,7	11,98	91,45	66,06	316
May1996	258,1	11,39	92,1	69,13	322
Jun1996	228,3	12,17	86,46	69,74	308

Jul1996	203,9	12,81	78,14	67,84	312
Aug1996	191,6	12,37	80,16	68,04	317
Sep1996	178,8	11,92	74,83	66,96	319
Oct1996	180,7	11,12	72,97	66,99	290
Nov1996	176,6	10,72	70,51	66,65	286
Dec1996	178,7	10,74	63,08	66,88	289
Jan1997	177,1	10,69	67,66	64,8	301
Feb1997	171,9	10,81	76,65	62,27	308
Mar1997	176,4	11,11	81,31	69,11	332
Apr1997	183,9	11,3	78,48	71,26	338
May1997	171,2	11,1	95,74	71,14	298
Jun1997	148,2	11,43	91,94	76,58	278
Jul1997	140,5	11,22	82,52	76,08	306
Aug1997	154,8	11,7	76,92	74,94	265
Sep1997	154,4	11,33	77,43	80,28	269
Oct1997	154,9	11,41	76,9	78,99	278
Nov1997	150,5	12,01	78,2	76,87	288
Dec1997	146,1	12,33	84,65	78,8	284
Jan1998	141	11,55	86,03	75,67	273
Feb1998	141,5	10,72	85,79	74,43	273
Mar1998	140,5	10,27	84,67	77,98	265
Apr1998	133	9,69	90,6	78,31	259
May1998	131	9,23	92,64	81,39	242
Jun1998	124,2	8,1	84,55	78,11	242
Jul1998	119,7	8,64	78,4	77,67	238
Aug1998	113	8,48	79,98	76,37	219
Sep1998	115,9	7,23	80,88	76,54	216
Oct1998	130,97	7,46	80,36	76,68	223
Nov1998	133,73	8,06	80,4	72,35	234
Dec1998	128,79	8,08	82,82	68,71	230
Jan1999	126,95	8,11	81,65	66,01	221
Feb1999	117,17	6,82	77,68	63,89	205
Mar1999	118,98	6,03	72,7	59,57	205
Apr1999	117,1	5,42	68,89	53,8	207
May1999	113,56	5,75	68,28	48,2	199
Jun1999	113,4	6,03	66,2	52,7	194
Jul1999	109,49	5,38	62,28	50,503	183
Aug1999	116,94	5,75	63,8	47,917	199
Sep1999	118,25	6,67	60,44	48,144	207
Oct1999	113,13	6,77	59,25	46,335	203
Nov1999	110,39	6,5	64,1	41,843	198
Dec1999	107,09	6,01	66,4	41,671	199
Jan2000	111,5	5,61	53,62	41,627	208
Feb2000	113,25	5,29	49,41	38,97	214
Mar2000	112,4	5,14	47,26	42,32	221
Apr2000	112	6,02	45,21	41,34	229
May2000	116,5	6,91	45,19	41,23	229

Jun2000	118	8,37	43,72	42,71	211
Jul2000	116	9,63	41,93	42,49	197
Aug2000	114	10,44	38,94	39,82	198
Sep2000	124	10,03	39,47	40,06	209
Oct2000	131	10,75	36,55	39,83	203
Nov2000	130,25	9,93	33,34	36,33	206
Dec2000	131,2	10,02	30,78	36,45	217
Jan2001	135,75	10,32	32,97	43,844	210
Feb2001	131,25	9,87	31,96	52,48	200
Mar2001	131	9,27	30,96	51	195
Apr2001	131,75	8,745	28,59	49,06	186
May2001	136	9,58	29,71	48,97	184
Jun2001	130	9,04	29,33	44,097	188
Jul2001	127,5	8,79	27,59	43,87	208
Aug2001	127	8,11	25,86	46,91	212
Sep2001	126,5	7,64	23,79	46,257	202
Oct2001	127,25	6,79	21,26	49,21	187
Nov2001	129,4	7,69	22,03	56,284	189
Dec2001	125,25	7,83	23,57	60,64	189
Jan2002	130	7,78	22,88	62,79	188
Feb2002	127	6,52	24,46	67,67	187
Mar2002	125,8	6,7	29,77	72,21	191
Apr2002	125,25	6,89	30,35	71,18	195
May2002	123,62	6,07	29,43	72,82	197
Jun2002	135,75	5,75	29,26	75,2368	200
Jul2002	151,2	6,39	29,31	84,76	218
Aug2002	165,22	6,29	28,74	88,927	227
Sep2002	191,8	6,96	33,31	98,156	236
Oct2002	194,36	7,51	34,44	100,031	229
Nov2002	182,76	7,83	39,38	82,28	243
Dec2002	164,67	7,95	38,68	91,751	241
Jan2003	154,72	8,19	42,75	99,1	244
Feb2003	153,83	9,01	42,35	101,57	243
Mar2003	144,83	8,23	38,36	90,48	241
Apr2003	143,25	7,75	38,68	87,44	251
May2003	146,21	7,17	38,9	79,367	250
Jun2003	132,72	6,73	35,33	71,63	239
Jul2003	133,78	6,84	36,71	70,78	230
Aug2003	155,13	6,83	37,92	71,28	237
Sep2003	149,82	5,98	38,76	74,01	264
Oct2003	151,05	5,96	37,32	67,22	310
Nov2003	169,33	6,08	36,05	68,54	327
Dec2003	169,9	6,28	37,59	73,81	332
Jan2004	170,07	5,81	41,32	73,677	350
Feb2004	164,12	5,84	39,1	71,175	368
Mar2004	169,04	6,46	38,61	68,241	413
Apr2004	171,89	6,58	38,02	65,198	358

May2004	164,28	6,34	38,04	64,128	316
Jun2004	157,54	6,93	41,09	63,865	295
Jul2004	152,41	7,91	36,44	70,7376	279
Aug2004	145,29	7,58	34,81	78,4374	265
Sep2004	155,52	7,69	35,1	69,8676	260
Oct2004	154,87	8,45	31,77	67,346	249
Nov2004	162,02	8,16	34,07	75,399	260
Dec2004	160,38	8,25	38,98	74,777	265
Jan2005	157,43	8,72	39,63	70,269	262
Feb2005	153,84	9,1	44,61	74,1455	261
Mar2005	157,56	8,88	50,7	79,72	290
Apr2005	147,91	8,59	53,32	71,94	283
May2005	149,27	8,59	58,66	68,56	283
Jun2005	150,08	9,05	62,96	69,81	306
Jul2005	148,69	9,64	60,57	67,54	298
Aug2005	158	9,93	55,6	67,22	274
Sep2005	169	10,31	50,07	68,23	263
Oct2005	174	11,12	50,84	65,89	257
Nov2005	166	11,38	54,72	65,04	256
Dec2005	168	13,31	58,79	68,839	264
Jan2006	171	15,75	66,46	71,568	257
Feb2006	185	17,95	65,5	70,0848	257
Mar2006	182	17,24	62,92	70,0417	257
Apr2006	186	17,49	64,45	70,4155	258
May2006	203	16,82	63,97	73,3243	266
Jun2006	205	15,4	64,14	72,8776	267
Jul2006	210	16,05	68,66	76,053	272
Aug2006	201	13,45	75,73	73,065	262
Sep2006	208	12,08	77,88	71,098	258
Oct2006	220	11,64	76,26	69,384	273
Nov2006	218	11,78	79,67	71,722	300
Dec2006	216	11,57	77,71	77,225	297
Jan2007	208	10,95	80,55	77,202	306
Feb2007	210	10,57	80,97	82,278	325
Mar2007	208	10,44	78,95	87,282	322
Apr2007	211	9,72	81,64	89,686	320
May2007	203	9,43	86,06	90,94	334
Jun2007	232	9,29	94,76	91,477	361
Jul2007	254	10,18	93,47	97,644	376
Aug2007	278	9,81	88,51	86,279	385
Sep2007	351	9,76	93,61	87,911	426
Oct2007	353	10	97,34	86,85	450
Nov2007	338	10,1	92,28	89,22	489
Dec2007	384	10,71	91,37	95,85	515
Jan2008	385	11,97	100,68	100,511	541
Feb2008	446	13,51	117,1	114,4465	572
Mar2008	454	13,2	122,44	121,1298	575

Apr2008	387	12,56	112,06	119,221	558
May2008	346	12,07	109,58	122,0012	575
Jun2008	369	12,13	112,16	137,067	622
Jul2008	346,195	14,23	115,09	133,978	634
Aug2008	342,726	14,61	113,48	127,482	556
Sep2008	309,7275	13,53	106,67	121,51	509
Oct2008	253,532	11,9	89,69	102,163	394
Nov2008	243,22	12,11	92,81	93,791	378
Dec2008	236,05	11,75	83,99	111,453	360
Jan2009	257,42	12,59	85,77	119,115	413
Feb2009	240	13,27	81,66	120,094	390
Mar2009	243	13,4	77,48	113,852	379
Apr2009	247	13,65	76,5	115,9025	414
May2009	269	16,04	77	112,526	465
Jun2009	261	16,43	75,88	122,4883	503
Jul2009	233	18,43	74,83	126,6156	463
Aug2009	216	22,435	75,04	134,1141	470
Sep2009	200	23,06	77,31	142,5601	421
Oct2009	209	22,64	76,68	152,9766	427
Nov2009	227	22,26	73,08	153,5041	440
Dec2009	221	23,53	74,68	158,6502	450
Jan2010	211	26,47	75,09	159,89	436
Feb2010	206	25,39	73,49	148,62	406
Mar2010	204	18,66	72,53	140,15	408
Apr2010	201	16,45	76,26	146,12	411
May2010	193	15,2	76,21	144,18	407
Jun2010	181	15,8	82,51	146,55	408
Jul2010	219	17,46	89,95	146,49	429
Aug2010	280	18,51	89,06	139,33	457
Sep2010	309	22,51	87,11	130,41	470
Oct2010	294	24,61	90,57	132,79	492

Πίνακας 6

Σαν δεύτερο βήμα, έγινε η κατασκευή του προγράμματος στη Matlab. Το πρόγραμμα που κατασκευάστηκε διαβάζει τα δεδομένα και χρησιμοποιεί τον αλγόριθμο της EMD για να τα αποσυνθέσει. Τα αποτελέσματα που δίνει ο EMD δίνονται σαν είσοδοι στο νευρωνικό δίκτυο που παράγει την τιμή πρόβλεψης. Τα νευρωνικά δίκτυα που χρησιμοποιήθηκαν για κάθε προϊόν έχουν διαφορετική αρχιτεκτονική. Για να επιλέξουμε το βέλτιστο νευρωνικό δίκτυο, ελέγξαμε τις τιμές σφάλματος που δίνει το δίκτυο για διαφορετικό πλήθος εποχών. Σε κάθε περίπτωση επιλέχτηκε το νευρωνικό δίκτυο που ελαχιστοποιεί το μέσο τετραγωνικό σφάλμα, το οποίο υπολογίζεται από τον παρακάτω τύπο:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(P_i - A_i)^2$$

Το μέσο τετραγωνικό σφάλμα MSE (Mean Square Error) προκύπτει παίρνοντας μια μέση τιμή του τετραγώνου του σφάλματος δηλαδή της διαφοράς της πραγματικής (A_i) από τη προβλεπόμενη (P_i) τιμή κατά μήκος των δυο δειγμάτων, είτε είναι το δείγμα εκπαίδευσης είτε το δείγμα δοκιμής.

Όπως αναφέρθηκε παραπάνω πρέπει να ορίζεται ποια δεδομένα θα αποτελούν το δείγμα εκπαίδευσης και ποια το δείγμα δοκιμής. Στην παρούσα εργασία ως δεδομένα εκπαίδευσης θεωρήθηκε το 80% των δεδομένων, δηλαδή 402 τιμές και ως δεδομένα δοκιμής ορίστηκε το υπόλοιπο 20% των δεδομένων, δηλαδή 99 τιμές.

Στους παρακάτω πίνακες παρουσιάζονται τα αποτελέσματα των νευρωνικών δικτύων για το διαφορετικό πλήθος εποχών:

<u>Σιτάρι (Wheat)</u>

Εποχές (epochs)	Μέσο τετραγωνικό σφάλμα(MSE)	Κρυμμένα επίπεδα και κόμβοι (Layers)	Συναρτήσεις ενεργοποίησης
175	43,57	[4,5,5]	Satlin, tansig, logsig
250	38,72	[5,9]	Tansig, satlin
263	39,25	[4,5,4,5]	Logsig,Logsig,Logsig,tansig
345	**38,57**	**[3,7]**	**Logsig, Logsig**

<u>Ζάχαρη (Sugar)</u>

Εποχές (epochs)	Μέσο τετραγωνικό σφάλμα(MSE)	Κρυμμένα επίπεδα και κόμβοι (Layers)	Συναρτήσεις ενεργοποίησης
174	**1,08**	**[5,5]**	**Logsig, tansig**
201	1,17	[4,4,4]	Tansig, logsig,tansig
221	1,52	[2,2]	Logsig,satlin
223	1,13	[6,6]	Tansig,tansig

<u>Καφές (Coffee)</u>

Εποχές (epochs)	Μέσο τετραγωνικό σφάλμα(MSE)	Κρυμμένα επίπεδα και κόμβοι (Layers)	Συναρτήσεις ενεργοποίησης
129	36,33	[5,5]	Logsig, logsig
137	56,51	[2,2,6,6]	Tansig, satlin, tansig, logsig
184	**31,78**	**[6,5]**	**Logsig, logsig**
244	32,12	[8,6]	Logsig, logsig

<u>Κακάο (Cocoa)</u>

Εποχές (epochs)	Μέσο τετραγωνικό σφάλμα(MSE)	Κρυμμένα επίπεδα και κόμβοι (Layers)	Συναρτήσεις ενεργοποίησης
117	**11,38**	**[6,6]**	**Logsig, tansig**
141	15,51	[6,5]	Tansig, logsig
156	11,89	[8,6,5]	Satlin, satlin, Logsig
172	14,78	[5,6,3,4,2]	Logsig, tansig, Logsig, Logsig, tansig

<u>Σόγια (Soybeans)</u>

Εποχές (epochs)	Μέσο τετραγωνικό σφάλμα(MSE)	Κρυμμένα επίπεδα και κόμβοι (Layers)	Συναρτήσεις ενεργοποίησης
164	133,52	[6,6]	Tansig, Logsig
193	130,41	[7,5,5]	Logsig, logsig, tansig
280	**106,36**	**[6,8]**	**Logsig, tansig**
346	123	[5,4]	Logsig, tansig

Όπου:
satlin: Piecewise Linear Function
logsig: λογιστική συνάρτηση
Tansig:συνάρτηση υπερβολικής εφαπτομένης

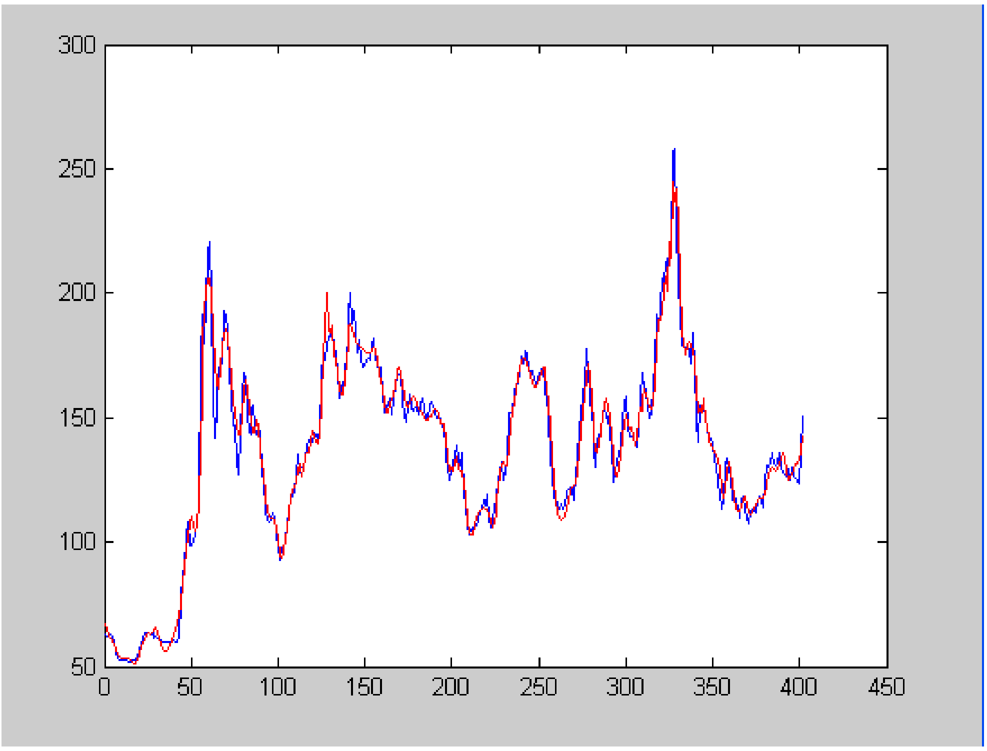
Όπως αναφέρθηκε επιλέχτηκε για κάθε προϊόν το πλήθος των εποχών που ελαχιστοποιεί το μέσο τετραγωνικό σφάλμα.

Παρακάτω παρουσιάζονται τα αποτελέσματα που δίνει το νευρωνικό δίκτυο που επιλέχτηκε σε κάθε περίπτωση, σε συνδυασμό με τον EMD. Στα διαγράμματα που θα ακολουθήσουν οι μπλε γραμμές αναπαριστούν τις πραγματικές τιμές των προϊόντων και οι κόκκινες γραμμές αναπαριστούν τις προβλεπόμενες τιμές που δίνει η μέθοδος. Το πρώτο διάγραμμα αφορά τα δεδομένα εκπαίδευσης και το δεύτερο τα δεδομένα δοκιμής.

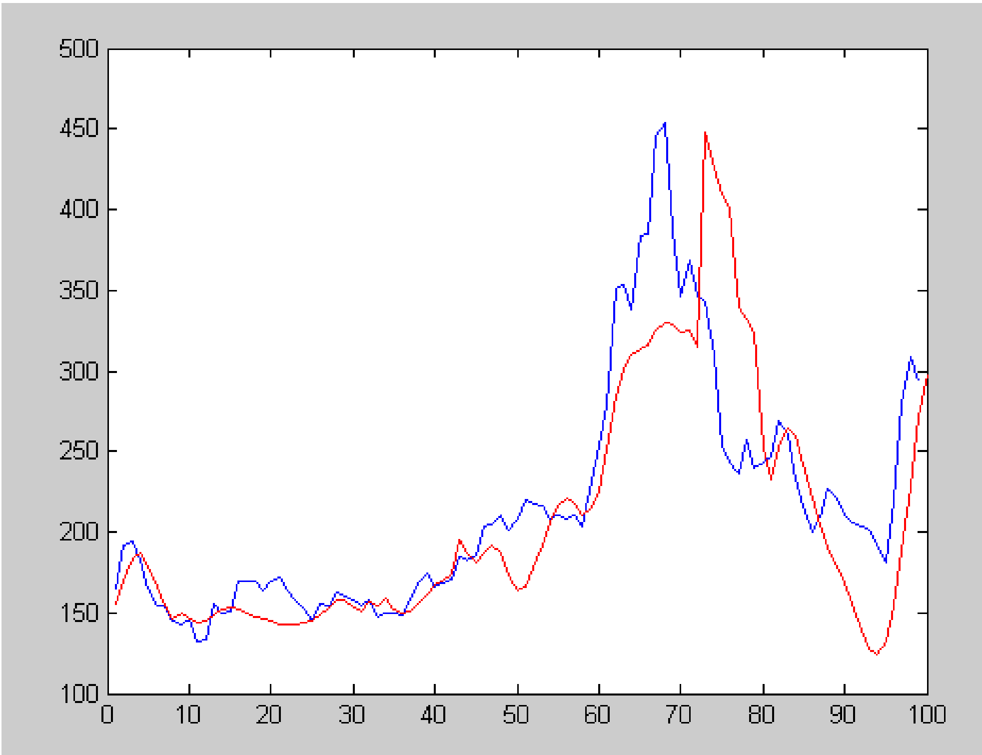
<u>Σιτάρι (Wheat)</u>

Επιλέχτηκε το νευρωνικό με τα παρακάτω στοιχεία:

- Πλήθος εποχών: 345
- Μέσο τετραφωνικό σφάλμα: 37,57
- Δύο κρυμμένα επίπεδα, με 3 και 7 κόμβους αντίστοιχα
- Η συνάρτηση ενεργοποίησης που χρησιμοποιήθηκε και στα δύο επίπεδα είναι η logsig



Διάγραμμα 1: Δεδομένα εκπαίδευσης

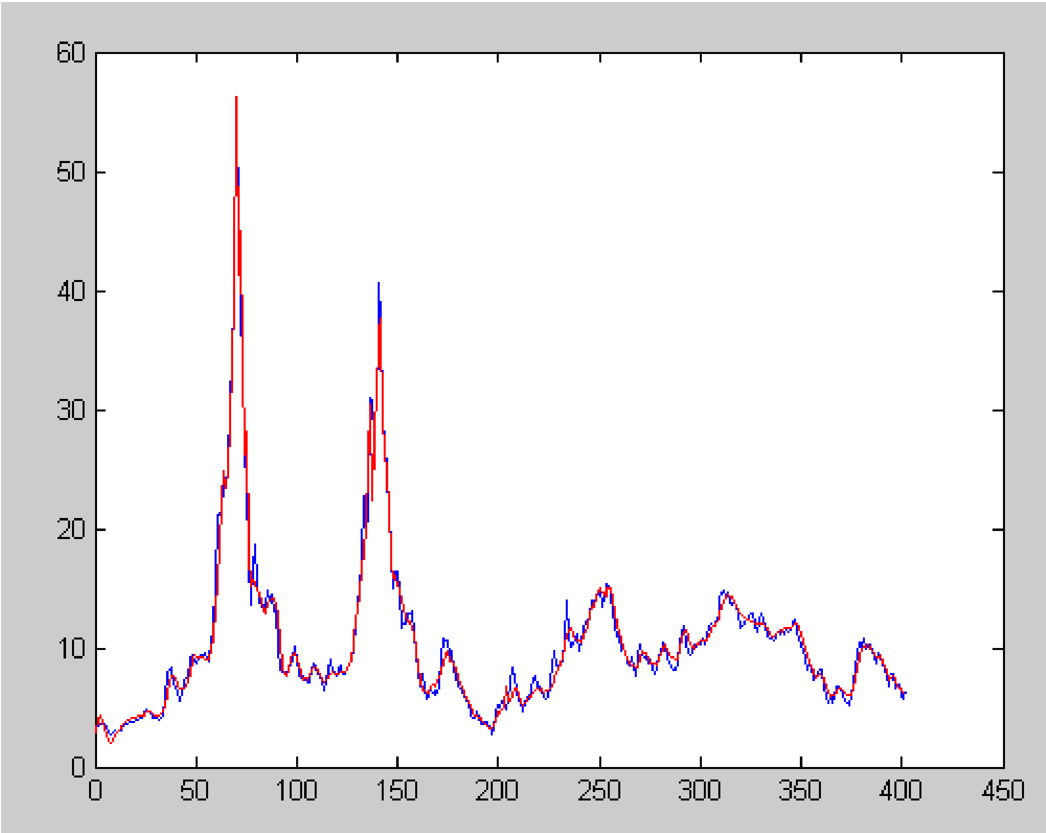


Διάγραμμα 2: Δεδομένα δοκιμής

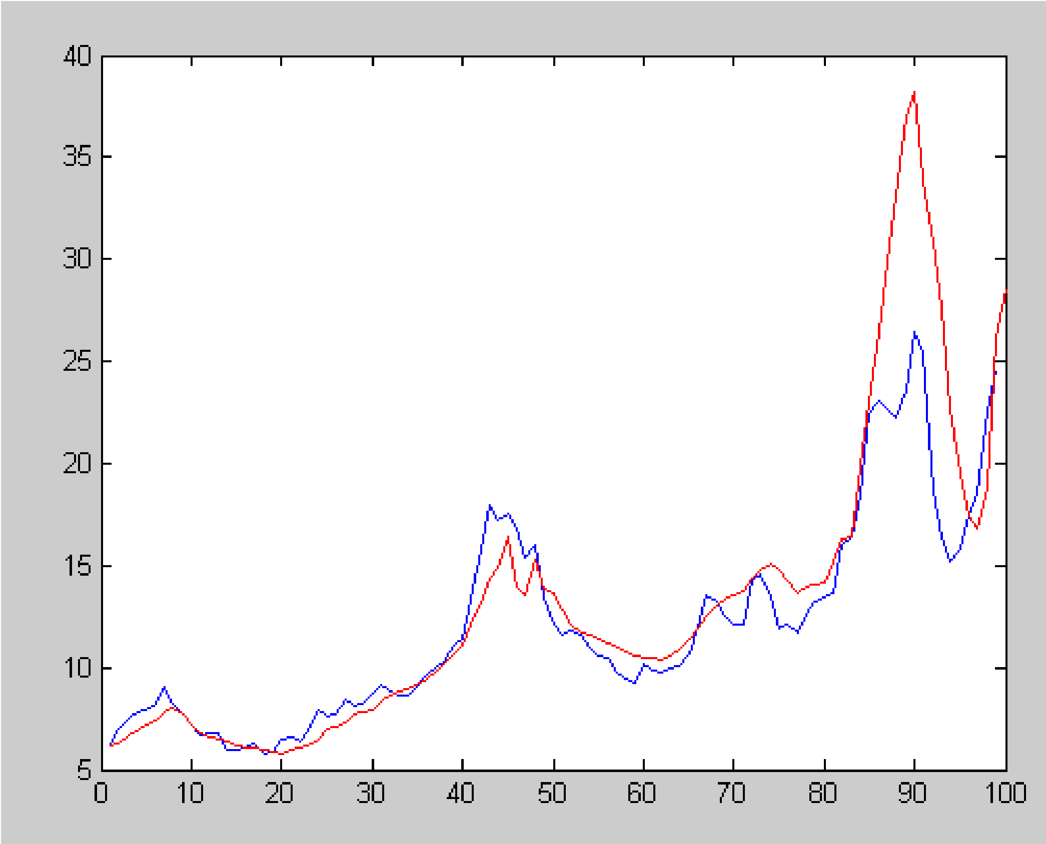
Ζάχαρη (Sugar)

Επιλέχτηκε το νευρωνικό με τα παρακάτω στοιχεία:

- Πλήθος εποχών: 174
- Μέσο τετραφωνικό σφάλμα: 1,08
- Δύο κρυμμένα επίπεδα, με 5 και 5 κόμβους αντίστοιχα
- Οι συναρτήσεις ενεργοποίησης που χρησιμοποιήθηκαν στα δύο επίπεδα είναι η logsig και η tansig



Διάγραμμα 3: Δεδομένα εκπαίδευσης

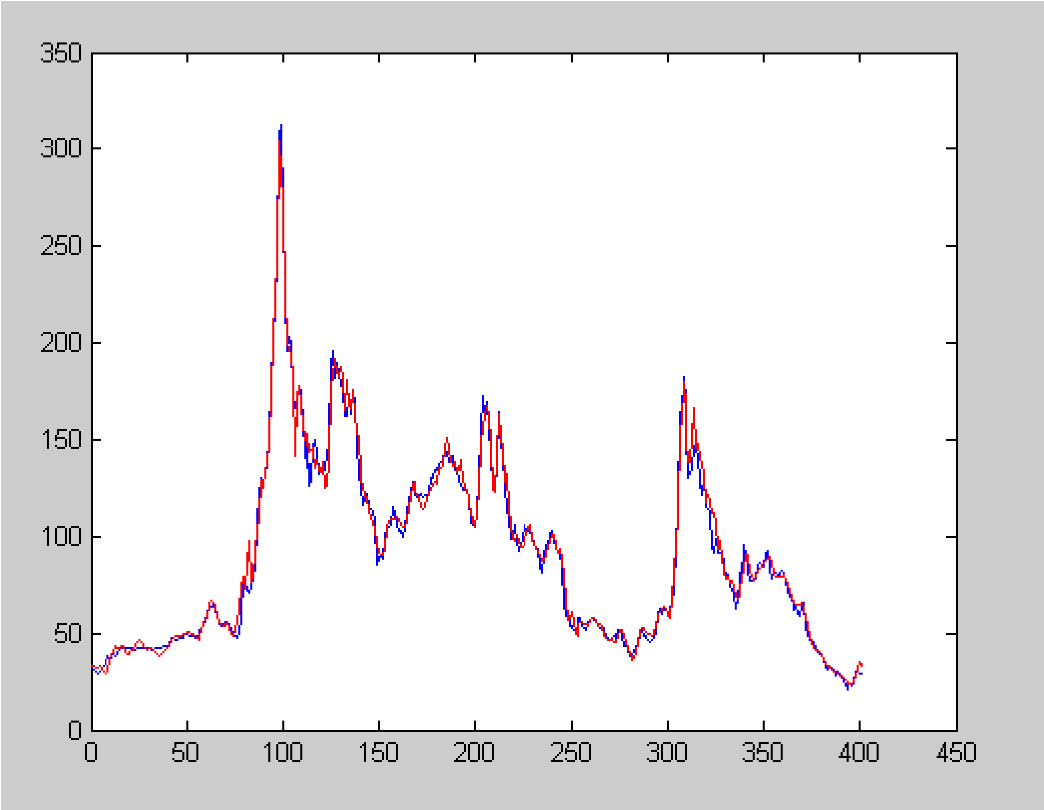


Διάγραμμα 4: Δεδομένα δοκιμής

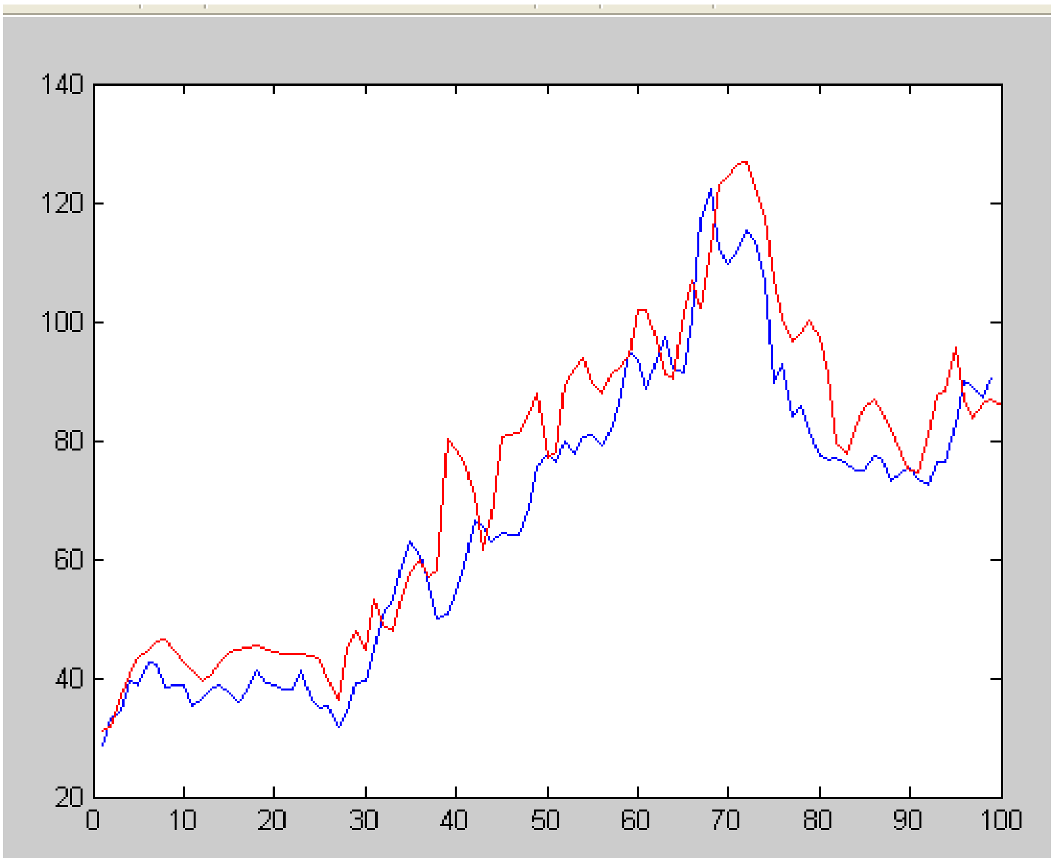
Καφές (Coffee)

Επιλέχτηκε το νευρωνικό με τα παρακάτω στοιχεία:

- Πλήθος εποχών: 184
- Μέσο τετραφωνικό σφάλμα: 31,78
- Δύο κρυμμένα επίπεδα, με 6 και 5 κόμβους αντίστοιχα
- Η συνάρτηση ενεργοποίησης που χρησιμοποιήθηκε και στα δύο επίπεδα είναι η logsig



<u>Διάγραμμα 5: Δεδομένα εκπαίδευσης</u>

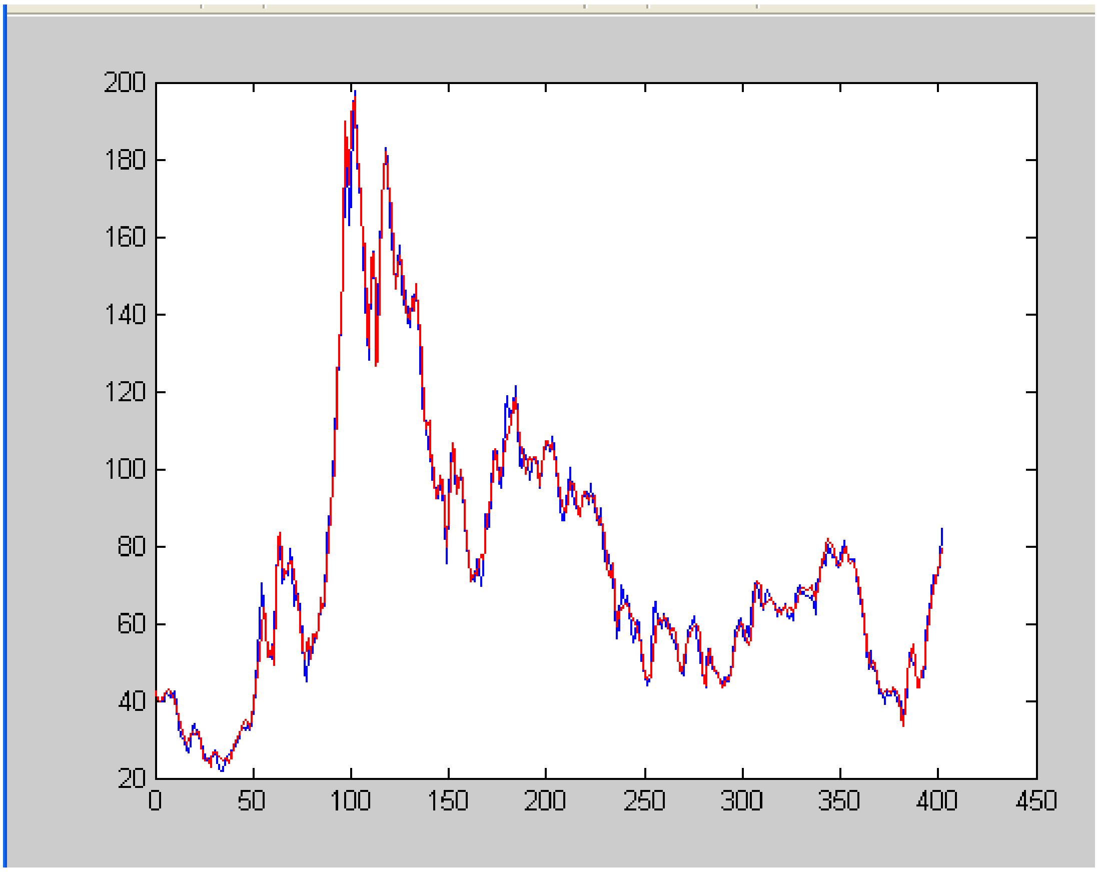


<u>Διάγραμμα 6: Δεδομένα δοκιμής</u>

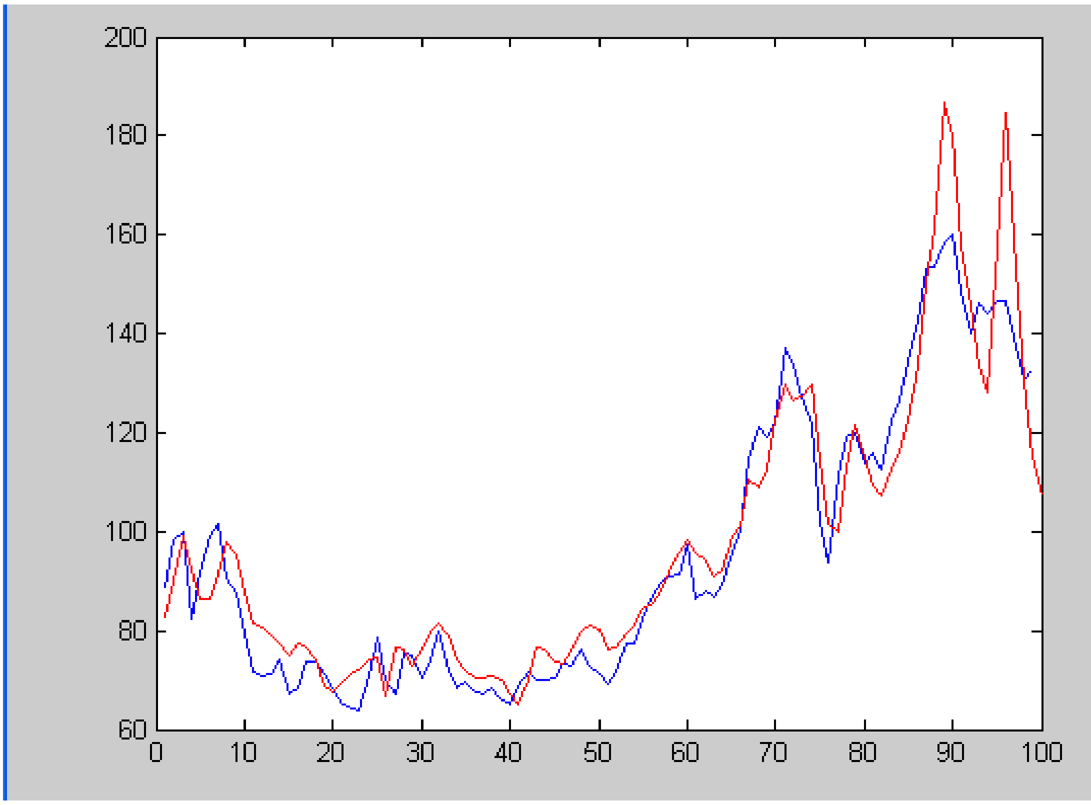
Κακάο (Cocoa)

Επιλέχτηκε το νευρωνικό με τα παρακάτω στοιχεία:

- Πλήθος εποχών: 117
- Μέσο τετραφωνικό σφάλμα: 11,38
- Δύο κρυμμένα επίπεδα, με 6 και 6 κόμβους αντίστοιχα
- Οι συναρτήσεις ενεργοποίησης που χρησιμοποιήθηκαν στα δύο επίπεδα είναι η logsig και η tansig



Διάγραμμα 7: Δεδομένα εκπαίδευσης

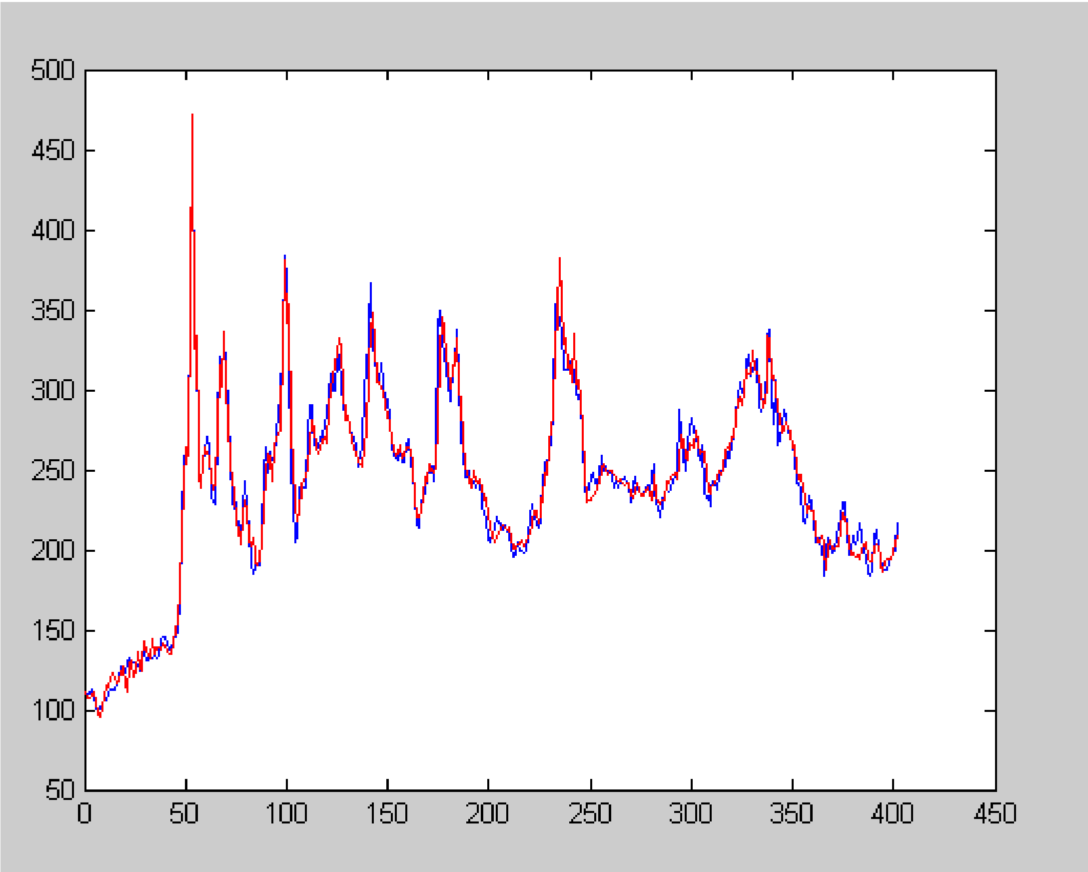


Διάγραμμα 8: Δεδομένα δοκιμής

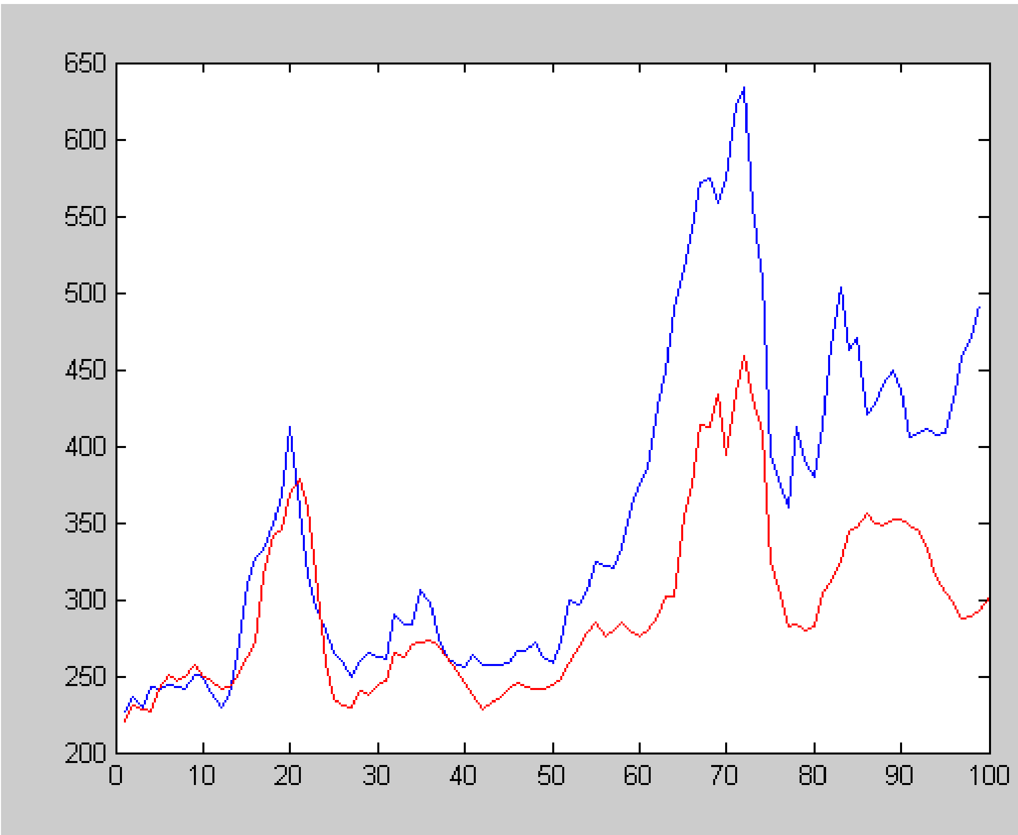
Σόγια (Soybeans)

Επιλέχτηκε το νευρωνικό με τα παρακάτω στοιχεία:

- Πλήθος εποχών: 280
- Μέσο τετραφωνικό σφάλμα: 106,36
- Δύο κρυμμένα επίπεδα, με 6 και 8 κόμβους αντίστοιχα
- Οι συναρτήσεις ενεργοποίησης που χρησιμοποιήθηκαν στα δύο επίπεδα είναι η logsig και η tansig



Διάγραμμα 9: Δεδομένα εκπαίδευσης



Διάγραμμα 10: Δεδομένα δοκιμής

Όπως μπορούμε να παρατηρήσουμε από τα αποτελέσματα που προέκυψαν, η μέθοδος EMD-BPN είναι κατάλληλη για την πρόβλεψη τιμών. Σύμφωνα με τα διαγράμματα εκπαίδευσης το μοντέλο έχει την ικανότητα να εκπαιδεύεται αρκετά καλά. Όσον αφορά τα διαγράμματα δοκιμής, οι προβλεπόμενες τιμές που δίνει η μέθοδος πλησιάζουν σε πολύ μεγάλο βαθμό τις πραγματικές τιμές. Σύμφωνα με τα διαγράμματα, μικρότερες αποκλίσεις μεταξύ πραγματικών και προβλεπόμενων τιμών παρουσιάζονται για την Ζάχαρη, για τον Καφέ και για το Κακάο, ενώ μεγαλύτερες αποκλίσεις τιμών παρουσιάζονται στο Σιτάρι και στη Σόγια.

Από τα δεδομένα του πίνακα 6, είναι εύκολο να παρατηρήσουμε τις μέγιστες και τις ελάχιστες τιμές για κάθε προϊόν. Σύμφωνα λοιπόν με τον πίνακα 6, προκύπτουν τα παρακάτω στοιχεία:

	Σιτάρι	**Ζάχαρη**	**Καφές**	**Κακάο**	**Σόγια**
Max	454	56,14	312,24	197,84	634
Min	51,8	2,74	21,26	21,79	99
Max-Min	402,2	53,4	290,98	176,05	535

Πίνακας 7

Όπως βλέπουμε, μεγαλύτερη απόκλιση μεταξύ μέγιστης και ελάχιστης τιμής δίνουν το Σιτάρι και η Σόγια. Δηλαδή τα προϊόντα που σύμφωνα με τα διαγράμματα παρουσιάζουν μεγαλύτερες αποκλίσεις μεταξύ πραγματικών και προβλεπόμενων τιμών. Μπορούμε λοιπόν να συμπεράνουμε ότι όσο πιο κοντά είναι μεταξύ τους οι τιμές που εκπαιδεύονται, δηλαδή όσο μικρότερη διασπορά έχουν, τόσο πιο σωστά θα εκπαιδευτούν. Αυτό σημαίνει πως οι προβλέψεις θα είναι πιο ακριβείς.

Αξιοσημείωτο είναι ότι όλα τα διαγράμματα δοκιμής τείνουν να εντοπίζουν τις μελλοντικές αυξομειώσεις των τιμών σε μεγάλο βαθμό, ακόμα κι αν δεν μπορούν να προβλέψουν τις ακριβείς τιμές.

ΚΕΦΑΛΑΙΟ 6: Συμπεράσματα

Τα νευρωνικά δίκτυα χρησιμοποιούνται όλο και περισσότερο στις μέρες μας, για προβλέψεις σε κάθε τομέα (οικονομικό, ιατρικό, βιομηχανικό κ.α.). Αυτό συμβαίνει διότι παρουσιάζουν μεγαλύτερη αξιοπιστία και ακριβέστερα αποτελέσματα από άλλες μεθόδους.

Η εμπειρική μέθοδος αποσύνθεσης (EMD) είναι μια σχετικά καινούρια μέθοδος που εφαρμόζεται ευρέως σε πολλά πεδία τα τελευταία χρόνια.

Ο συνδυασμός των μεθόδων EMD και BPN οδηγεί σε μια μέθοδο πρόβλεψης που μπορεί να δώσει πολύ ακριβή αποτελέσματα. Ο αλγόριθμος της EMD αποσυνθέτει τα περίπλοκα αρχικά δεδομένα, σε τμήματα με απλούστερη συχνότητα (IMFs) και κάποιο υπόλειμμα, που χρησιμοποιούνται σαν είσοδοι για το νευρωνικό δίκτυο. Εφόσον οι είσοδοι έχουν απλούστερη συχνότητα, το νευρωνικό δίκτυο μπορεί να τις επεξεργαστεί ευκολότερα και να καταλήξει σε ακριβέστερα αποτελέσματα, σε σχέση με τα αποτελέσματα που θα έδινε χωρίς την αποσύνθεση των αρχικών δεδομένων.

Στην παρούσα εργασία έγινε μια προσπάθεια αξιολόγησης της μεθόδου EMD-BPN στην πρόβλεψη των τιμών κάποιων προϊόντων. Σύμφωνα με τα αποτελέσματα της εργασίας, η μέθοδος είναι κατάλληλη για την πρόβλεψη τιμών. Πιο συγκεκριμένα όμως, η μέθοδος καταλήγει σε ακριβέστερα αποτελέσματα όταν δίνονται αρκετά δεδομένα, ώστε να μπορέσει να εκπαιδευτεί σωστά το νευρωνικό δίκτυο. Επίσης, όσο μικρότερες αποκλίσεις παρουσιάζουν μεταξύ τους οι τιμές των δεδομένων τόσο πιο ακριβή αποτελέσματα θα παράγει η μέθοδος.

Βιβλιογραφία

- Anna Linderhed, 2004 ,Image Compression based on Empirical Mode Decomposition,http://aquador.vovve.net/IEMD/SSBA-04_corrected.pdf

- Στέλιος Χόης, (2005), Νευρωνικά Δίκτυα στον Αυτόματο Έλεγχο, http://utopia.duth.gr/~schois/Ergasia_Neuronika.pdf

- http://www.google.gr/imgres?q=empirical+mode+decomposition&hl=el&biw=1360&bih=601&tbm=isch&tbnid=4mudYp_X05YySM:&imgrefurl=http://www.music.mcgill.ca/~doug/code.html&docid=GvGBi5SafRBz8M&imgurl=http://www.music.mcgill.ca/~doug/emd_googlecode.png&w=499&h=251&ei=wsbWTvSiAcPKsgbS__3vCw&zoom=1&iact=rc&dur=223&sig=114605110381708280056&page=4&tbnh=79&tbnw=157&start=76&ndsp=25&ved=1t:429,r:14,s:76&tx=20&ty=20

- Καβαλιεράτου,(2010),νευρωνικάδίκτυα,http://www.icsd.aegean.gr/lecturers/kavallieratou/PattRec_files/4other_class.pdf

- Ζερβουδάκης Ιωάννης, 2010, "Πρόβλεψη εξωτερικής θερμοκρασίας με χρήση νευρωνικών δικτύων"

- http://el.wikipedia.org/wiki/%CE%9D%CE%B5%CF%85%CF%81%CF%89%CE%BD%CE%B9%CE%BA%CF%8C_%CE%B4%CE%AF%CE%BA%CF%84%CF%85%CE%BF

- Ι. Βλαχάβας , Π. Κεφαλάς, Ν. Βασιλειάδης, Φ. Κόκκορας, Η. Σακελαρίου, (2005), Τεχνητά νευρωνικά δίκτυα (Artificial neural nets), http://aibook.csd.auth.gr/include/slides/Chap19.pdf

- Yongxiang Huang a,b, Franηois G. Schmitt b,*, Zhiming. Lu a, Yulu Liu a, 2009, Analysis of daily river flow fluctuations using empirical mode decomposition and arbitrary order Hilbert spectral analysis, Journal of Hydrology

- Ευτύχιος Πρωτοπαπαδάκης, 2009 , «Πρόβλεψη τάσης των τιμών των μετοχών σε περιόδους χρηματιστηριακών κρίσεων»

- B.L. Barnhart n, W.E.Eichinger, 2011, Empirical Mode Decomposition applied to solar irradiance, global temperature, sunspot number, andCO2 concentration data, Journal of Atmospheric and Solar-Terrestrial Physics

- Huang et al., 1998. Flandrin, Rilling and Goncalves, 2004. Huang, 2001.Wu and Hu, 2006

- Μιχάλης Ψωματάκης ,2009, Πρόβλεψη τιμών χρηματιστηριακών αγορών με τεχνική ανάλυση και νευρωνικά δίκτυα

- Donghoh Kim and Hee-Seok Oh, 2009, EMD: A Package for Empirical Mode Decomposition and Hilbert Spectrum

- Παρασύρη Μαρία, 2010, Πρόβλεψη γαλακτοπαραγωγής με χρήση νευροασαφούς μοντέλου(ANFIS)

- http://www.google.gr/imgres?q=empirical+mode+decomposition&hl=el&biw=1360&bih=601&tbm=isch&tbnid=DaSDaDjbXhfpAM:&imgrefurl=http://www.worldclimatereport.com/index.php/2007/03/16/the-coming-globalcooling/&docid=TjwCgAx6tmD3pM&imgurl=http://www.worldclimatereport.com/wpimages/global_cooling.JPG&w=565&h=434&ei=wsbWTvSiAcPKsgbS__3vCw&zoom=1&iact=rc&dur=359&sig=114605110381708280056&page=1&tbnh=128&tbnw=163&start=0&ndsp=23&ved=1t:429,r:21,s:0&tx=42&ty=90

- Γ.Α.Ευθήμερος 2010, «Ανάπτυξη ενός συστήματος τεχνητής νοημοσύνης ενεργού ελέγχου δονήσεων και θορύβου με τη χρήση ενός τεχνητού νευρωνικού δικτύου και ενός γενετικού αλγορίθμου» http://nemertes.lis.upatras.gr/jspui/bitstream/10889/4562/1/Efth_01-10_2.pdf

- http://nemertes.lis.upatras.gr/dspace/handle/123456789/1098?mode=full&submit_simple=%CE%95%CE%BC%CF%86%CE%B1%CE%BD%CE%AF%CF%83%CF%84%CE%B5+%CF%84%CE%B7%CE%BD+%CF%80%CE%BB%CE%AE%CF%81%CE%B7+%CE%B5%CE%B3%CE%B3%CF%81%CE%B1%CF%86%CE%AE

- http://nemertes.lis.upatras.gr/dspace/handle/123456789/3545

- Chauvin & Rumelhart, 1995

- Yu and Schwartz 2006

- Burger, Dohnal, Kathrada, & Law, 2001- Cho, 2003- Jeng &Fesenmaier, 1996- Law, 1998, 2000, 2001- Law & Au, 1999- Mohsin & Ryan, 1999-Pattie & Snyder, 1996- Sirakaya, Delen & Choi, 2005- Tsaur, Chiu & Huang, 2002-Uysal & El Roubi, 1999- Wang, 2004- Palmer, Jose Montano, & Sese, 2006
- Chauvin & Rumelhart, 1995.Haykin, 1999. McNelis, 2004
- Huang, Shen and Long, 1999
- Yu, Cheng and Yang, 2005. Cheng, Yu and Yang, 2004
- Vincent, Hu and Hou, 1999