

Time Lagged Feedforward Network (TLFN)

Νευρωνικό Δίκτυο Προώθησης με Χρονική Υστέρηση

Καπαγερίδης Ιωάννης

Γενικά

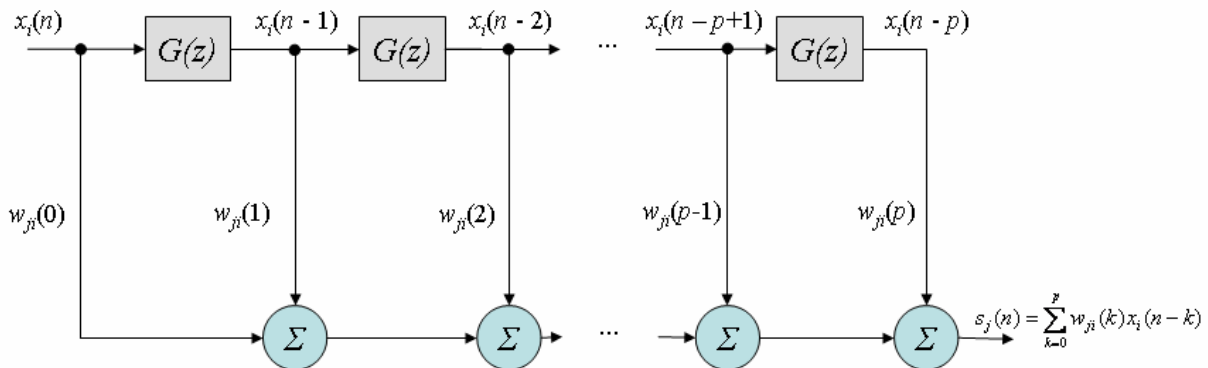
Τα Νευρωνικά Δίκτυα Προώθησης με Χρονική Υστέρηση (Time-Lagged Feedforward Networks - TLFNs) αποτελούν μία παραλλαγή των κλασικών νευρωνικών δικτύων προώθησης, ειδικά σχεδιασμένα για τη μοντελοποίηση χρονοσειρών και δεδομένων με χρονικές εξαρτήσεις. Χρησιμοποιούν δεδομένα με χρονική υστέρηση ως είσοδο για να ενσωματώσουν πληροφορίες από το παρελθόν. Είναι κατάλληλα για προβλήματα χρονοσειρών όπου η χρονική εξάρτηση είναι περιορισμένη σε σύντομα χρονικά διαστήματα (πχ. πρόβλεψη καιρού). Μπορούν να βοηθήσουν στην αναγνώριση χρονικών μοτίβων σε δεδομένα χρονοσειρών, δίνοντας απαντήσεις που δεν βασίζονται μόνο στις τρέχουσες τιμές εισόδου αλλά και σε προηγούμενες (Haykin, 1999).

Αρχιτεκτονική - Λειτουργία

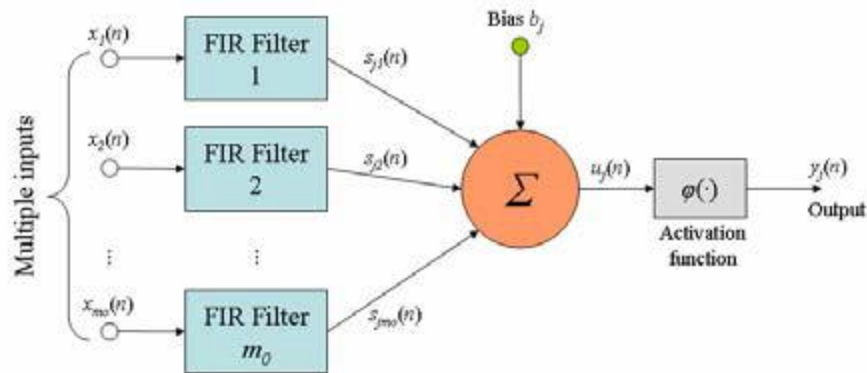
Τα δίκτυα TLFN βασίζονται στη χρήση ενός χωροχρονικού μοντέλου νευρώνα, δηλαδή ενός φίλτρου νευρώνα με πολλαπλές εισόδους ή κατανεμημένου νευρωνικού φίλτρου. Αυτό το μοντέλο χρησιμοποιεί **φίλτρα πεπερασμένης διάρκειας κρουστικής απόκρισης** (finite-duration impulse response filter - FIR) ως συνάψεις. Έτσι, το νευρωνικό φίλτρο με πολλαπλές εισόδους παρέχει ένα ισχυρό λειτουργικό μπλοκ για τη χωροχρονική επεξεργασία σήματος, κατασκευασμένο γύρω από έναν μόνο νευρώνα. Κάθε νευρώνας μνήμης στο δίκτυο TLFN επεξεργάζεται αποτελεσματικά πληροφορίες με την πάροδο του χρόνου, δουλεύοντας με τις προβολές των ενεργοποιήσεων των νευρώνων του προηγούμενου επιπέδου στον τοπικό γραμμικό χώρο μνήμης του (Principe et al. 2000). Το μέγεθος του κάθε χώρου μνήμης (δηλαδή, ο αριθμός των βάσεων) καθορίζεται από τον αριθμό των χρονικών καταγραφών μνήμης (memory taps).

Το κατανεμημένο νευρωνικό φίλτρο βασίζεται στην υπολογιστική ισχύ του φίλτρου πεπερασμένης διάρκειας κρουστικής απόκρισης (FIR) τάξης p , όπως φαίνεται στο Σχήμα 1. Το φίλτρο FIR αποτελεί ένα από τα βασικά δομικά στοιχεία στην ψηφιακή επεξεργασία σήματος (Haykin and Van Veen, 1998; Oppenheim and Schaffer, 1989). Η υπολογιστική ισχύς του φίλτρου FIR επεκτείνεται με τη χρήση πολλαπλών εισόδων, αριθμού m_0 , όπως απεικονίζεται στο Σχήμα 2. Το χωροχρονικό μοντέλο του κατανεμημένου νευρωνικού φίλτρου αναφέρεται επίσης ως φίλτρο νευρώνα με πολλαπλές εισόδους.

Ο νευρώνας διαθέτει m_0 κύριες συνάψεις, καθεμία από τις οποίες αποτελείται από ένα γραμμικό διακριτό φίλτρο χρόνου που υλοποιείται με τη μορφή φίλτρου FIR τάξης p . Οι κύριες συνάψεις αντιπροσωπεύουν τη χωρική διάσταση της επεξεργασίας σήματος. Κάθε κύρια σύναψη διαθέτει $(p + 1)$ δευτερεύουσες συνάψεις, οι οποίες συνδέονται με την αντίστοιχη είσοδο της και με τις χρονικές καταγραφές μνήμης (memory taps) του φίλτρου FIR, λαμβάνοντας έτσι υπόψη τη χρονική διάσταση της επεξεργασίας σήματος (Σχήμα 3).



Σχήμα 1: Φίλτρο πεπερασμένης διάρκειας κρουστικής απόκρισης (FIR).



Σχήμα 2: Καταμεμημένο (με πολλαπλές εισόδους) νευρωνικό φίλτρο.

Η χωροχρονική επεξεργασία που πραγματοποιείται από το νευρωνικό φίλτρο στο Σχήμα 4 μπορεί να εκφραστεί μαθηματικά με βάση την έξοδό του, $y_j(n)$, ως εξής:

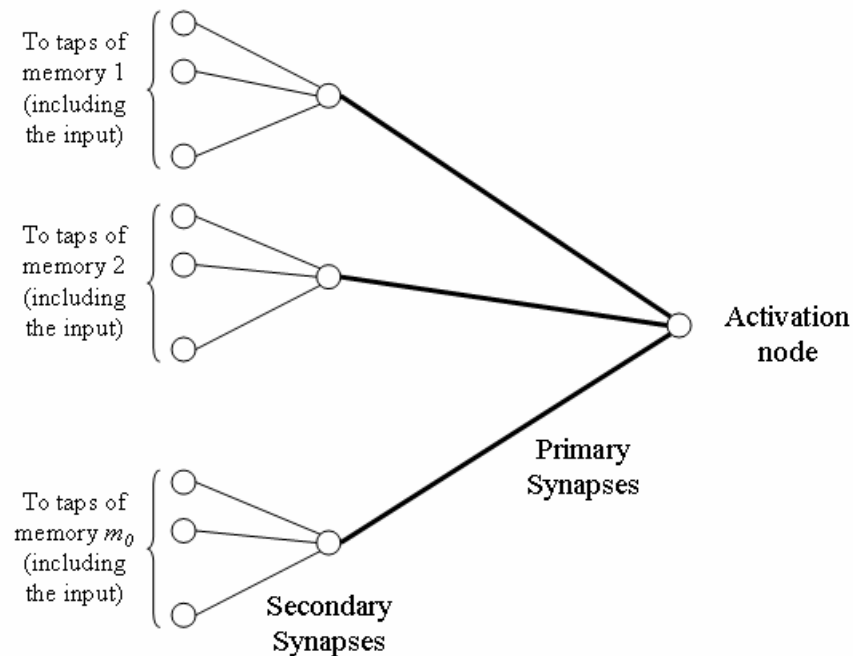
$$y_j(n) = \varphi \left(\sum_{i=1}^{m_0} \sum_{l=0}^p w_{ji}(l) x_i(n-l) + b_j \right)$$

Όπου $w_{ji}(l)$ είναι το βάρος της l δευτερεύουσας σύναψης που ανήκει στην i κύρια σύναψη, $x_i(n)$ είναι η είσοδος στην i κύρια σύναψη τη χρονική στιγμή n , και b_j είναι η μεροληψία (bias) που εφαρμόζεται στον νευρώνα. Η συνολική δομή του καταμεμημένου δικτύου TLFN παρουσιάζεται στο Σχήμα 4.

Εκπαίδευση

Για την εκπαίδευση ενός καταμεμημένου δικτύου TLFN απαιτείται ένας αλγόριθμος επιβλεπόμενης μάθησης, όπου η πραγματική απόκριση κάθε νευρώνα στο στρώμα εξόδου συγκρίνεται με την επιθυμητή απόκριση σε κάθε χρονική στιγμή. Ο αλγόριθμος χρονικής οπισθοδιάδοσης (temporal back

propagation) χρησιμοποιείται συχνά για την εκπαίδευση των δικτύων TLFN, όπως περιγράφεται από τον Haykin (1999). Ο αλγόριθμος αυτός προσαρμόζει τα βάρη ενός κατανεμημένου νευρωνικού δικτύου (όπως το TLFN) με βάση τις αποκρίσεις του δικτύου σε διαφορετικές χρονικές στιγμές, συγκρίνοντας τις προβλέψεις του με τις επιθυμητές αποκρίσεις. Οι προσαρμογές αυτές πραγματοποιούνται λαμβάνοντας υπόψη τόσο την χωρική όσο και τη χρονική πληροφορία, ώστε να ενισχύεται η ικανότητα του δικτύου να επεξεργάζεται δυναμικά δεδομένα.



Σχήμα 3: Συναπτική δομή ενός κατανεμημένου νευρωνικού φίλτρου.

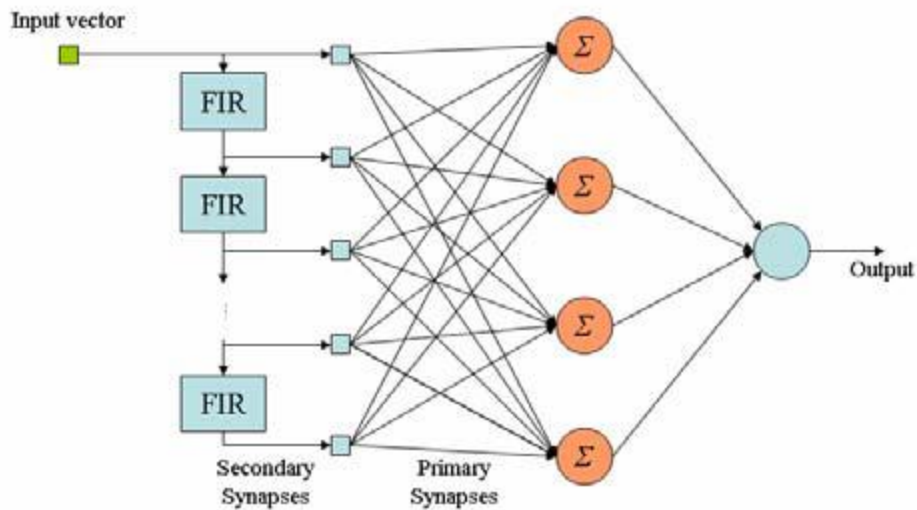
Η λειτουργία του αλγορίθμου μπορεί να περιγραφεί σε τρία βασικά στάδια:

1. Προώθηση Σήματος (Forward Pass)

Αρχικά, το δίκτυο υπολογίζει την έξοδό του (ή τις εξόδους του) για κάθε χρονική στιγμή, λαμβάνοντας υπόψη τόσο:

- τα **δεδομένα εισόδου** σε κάθε χρονική στιγμή, όσο και
- τις **εσωτερικές καταστάσεις του δικτύου**, που προκύπτουν από προηγούμενες χρονικές στιγμές (αυτό είναι χαρακτηριστικό στα χωροχρονικά μοντέλα).

Κατά την προώθηση σήματος, το δίκτυο χρησιμοποιεί τις **πρωτεύουσες και δευτερεύουσες συνάψεις** (π.χ. τα φίλτρα FIR) για να επεξεργαστεί τόσο τη χωρική όσο και τη χρονική πληροφορία. Έτσι, η έξοδος ενός νευρώνα σε χρόνο n , δηλαδή $y_j(n)$, υπολογίζεται ως συνάρτηση των εισόδων του $x_i(n)$, των βαρών των συνάψεων w_{ji} , και της μνήμης του (memory taps).



Σχήμα 4: Αρχιτεκτονική δικτύου TLFN.

2. Υπολογισμός του Σφάλματος (Error Calculation)

Σε κάθε χρονική στιγμή n , συγκρίνεται η **πραγματική έξοδος του δικτύου** $y_j(n)$ με την **επιθυμητή έξοδο** $d_j(n)$. Το σφάλμα εξόδου $e_j(n)$ για τον κάθε νευρώνα υπολογίζεται ως:

$$e_j(n) = d_j(n) - y_j(n)$$

Αυτό το σφάλμα περιγράφει πόσο απέχει η πρόβλεψη του δικτύου από την επιθυμητή απάντηση σε χρόνο n .

3. Οπισθοδιάδοση Σφάλματος (Backward Pass)

Η καινοτομία του χρονικού αλγορίθμου είναι ότι το σφάλμα δεν οπισθοδιαδίδεται μόνο μέσω των συνάψεων (όπως στην κλασική οπισθοδιάδοση), αλλά και **μέσα στον χρόνο**. Δηλαδή:

- Το σφάλμα που υπολογίζεται σε μια χρονική στιγμή επηρεάζει την προσαρμογή των βαρών για **προηγούμενες χρονικές στιγμές**.
- Λαμβάνονται υπόψη τα **memory taps** (δηλαδή οι χρονικές καθυστερήσεις) που εισάγονται από τα φίλτρα FIR, ώστε το σφάλμα να διαχέεται σωστά στον χρόνο.

Για να γίνει αυτό, υπολογίζονται οι **παράγωγοι μερικών συναρτήσεων κόστους** (π.χ. το Mean Squared Error - MSE) ως προς τα βάρη των συνάψεων, όχι μόνο για την τρέχουσα χρονική στιγμή, αλλά για όλες τις προηγούμενες στιγμές όπου το συγκεκριμένο βάρος επηρέασε την έξοδο του δικτύου. Η εξίσωση ενημέρωσης των βαρών για το βάρος $w_{ji}(l)$ έχει γενικά τη μορφή:

$$w_{ji}(l) \leftarrow w_{ji}(l) + \eta \cdot \frac{\partial E}{\partial w_{ji}(l)}$$

όπου:

- η είναι ο ρυθμός μάθησης (learning rate),
- E είναι η συνάρτηση κόστους (π.χ. το συνολικό σφάλμα στο χρόνο),

$$\frac{\partial E}{\partial w_{ji}(l)}$$

- είναι η παράγωγος του κόστους ως προς το βάρος, η οποία υπολογίζεται λαμβάνοντας υπόψη τις χωροχρονικές εξαρτήσεις.

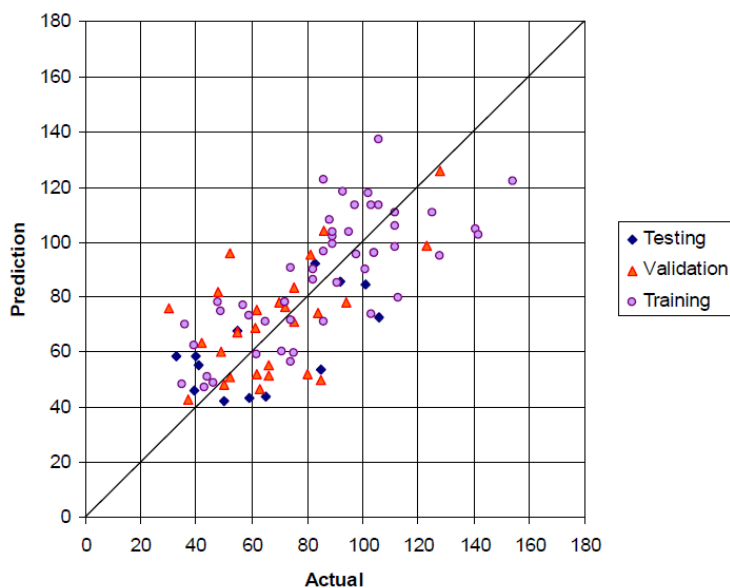
Παράδειγμα Εφαρμογής

Οι συγκεντρώσεις των αιωρούμενων σωματιδίων στην ατμόσφαιρα έχει βρεθεί ότι συνδέονται με μια ευρεία γκάμα επιπτώσεων στην ανθρώπινη υγεία. Οι συγκεντρώσεις PM10 χρησιμοποιούνται συνήθως ως τυπικό μέτρο για τη ρύπανση του αέρα. Η αύξηση των επιπέδων PM10 έχει συσχετιστεί με αυξήσεις στη θνησιμότητα και στις καρδιοαναπνευστικές νοσηλείες. Ως εκ τούτου, η πρόβλεψη των επιπέδων συγκέντρωσης σε συγκεκριμένα περιβάλλοντα είναι μεγάλης σημασίας, ειδικά σε αστικές και βιομηχανοποιημένες περιοχές.

Οι Karageridis και Triantafyllou (2004) σε εργασία τους ανέπτυξαν ένα προσαρμοστικό σύστημα βασισμένο σε TLFN με γενετικά βελτιστοποιημένη τοπολογία και παραμέτρους εκμάθησης, το οποίο επιτρέπει την πρόβλεψη της μέγιστης κινούμενης μέσης τιμής 24 ωρών της συγκέντρωσης PM10. Το δίκτυο εκπαιδεύεται και αξιολογείται με ωριαία δεδομένα που συλλέχθηκαν από δύο σταθμούς παρακολούθησης ατμοσφαιρικών ρύπων, σε μια αστική και μια κοντινή βιομηχανική περιοχή στη βόρεια Ελλάδα. Η αρχική μελέτη που παρουσιάζεται σε αυτή την εργασία αφορά ένα μικρό υποσύνολο των διαθέσιμων δεδομένων που χρησιμοποιήθηκαν για να επικυρωθεί η προσέγγιση και η επιλεγμένη αρχιτεκτονική ANN.

Εφαρμόστηκε γενετικός αλγόριθμος για την επιλογή του βέλτιστου αριθμού κρυφών κόμβων, των παραμέτρων εκμάθησης και της γενικής τοπολογίας του δικτύου. Το τελικό βέλτιστο δίκτυο περιείχε 10 κρυφές μονάδες, βήμα εκμάθησης 0.55 και ορμή 0.7. Τα δεδομένα προέρχονται από δύο σταθμούς παρακολούθησης ρύπων στη βόρεια Ελλάδα (Κοζάνη), περιλαμβάνοντας μετρήσεις PM10 και μετεωρολογικές παραμέτρους (υγρασία, θερμοκρασία, ταχύτητα ανέμου). Εξετάστηκε μια περίοδος τριών μηνών με ωριαίες μετρήσεις και υπολογίστηκε η μέγιστη κινούμενη μέση τιμή 24 ωρών του PM10 για την επόμενη μέρα.

Το βέλτιστο TLFN δίκτυο πέτυχε υψηλή ακρίβεια πρόβλεψης, με μέσο τετραγωνικό σφάλμα (MSE) 0.081 στο σετ εκπαίδευσης και 0.09 στο σετ επικύρωσης. Κατά τη δοκιμή του δικτύου σε δεδομένα που δεν είχαν χρησιμοποιηθεί στα στάδια ανάπτυξης, το μοντέλο πέτυχε MSE 0.082, δείχνοντας αξιόπιστες δυνατότητες γενίκευσης. Το TLFN αποδείχθηκε ιδιαίτερα αποτελεσματικό στην πρόβλεψη της μέγιστης συγκέντρωσης PM10 για την επόμενη ημέρα (Σχήμα 5). Η μελέτη έδειξε την καταλληλότητα του συγκεκριμένου μοντέλου για την αντιμετώπιση προβλημάτων χωροχρονικής ανάλυσης, ειδικά σε περιβάλλοντα που υπόκεινται σε μεταβαλλόμενες συνθήκες ρύπανσης.



Σχήμα 5: Διάγραμμα διασποράς πραγματικών τιμών και προβλέψεων του TLFN μέγιστων 24ωρων συγκεντρώσεων PM10 στα τρία τμήματα του σετ δεδομένων.

Πρόταση Εφαρμογής

Τα TLFN θα μπορούσαν να χρησιμοποιηθούν σε εργοστάσια εμπλουτισμού μεταλλευμάτων για τη δυναμική πρόβλεψη και βελτιστοποίηση της απόδοσης των διεργασιών εμπλουτισμού, όπως ο διαχωρισμός, η ανάκτηση μετάλλων και η ποιότητα του τελικού συμπυκνώματος. Χρησιμοποιώντας δεδομένα εισόδου από αισθητήρες, τα TLFN μπορούν να προβλέψουν την ποιότητα του προϊόντος και να εντοπίσουν αποκλίσεις από τις βέλτιστες συνθήκες λειτουργίας, επιτρέποντας την έγκαιρη προσαρμογή παραμέτρων για μέγιστη αποδοτικότητα. Επιπλέον, μπορούν να συμβάλουν στη διαχείριση της ανακύκλωσης υπολειμμάτων (tailings), στη μείωση της κατανάλωσης ενέργειας και πόρων, καθώς και στην πρόβλεψη φθοράς εξοπλισμού, υποστηρίζοντας έτσι πιο αποδοτική και βιώσιμη λειτουργία του εργοστασίου.

Βιβλιογραφία

Haykin, S., and Van Veen, B., *Signals and Systems*. Wiley, 1998.

Haykin, S., *Neural networks – a comprehensive foundation* (2nd edition). Prentice Hall, 1999.

Kapageridis, I., and Triantafyllou, A., *A Genetically Optimised Neural Network for Prediction of Maximum Hourly PM10 Concentration*. In: 12th International Conference on Modelling, Monitoring and Management of Air Pollution (Air Pollution 2004), Wessex Institute of Technology, Rhodes 2004.

Oppenheim, A.V., and Schafer, R.W., *Discrete-time signal processing*. Englewood Cliffs, Prentice Hall, 1989.

Principe, J.C., Euliano, N.R., and Lefebvre, W.C., *Neural and adaptive systems – fundamentals through simulations*. John Wiley & Sons, 2000.