



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΩΝ ΔΙΑΤΑΞΕΩΝ & ΣΥΣΤΗΜΑΤΩΝ
ΑΠΟΦΑΣΕΩΝ

Βελτίωση της Μεθόδου Θήτα με χρήση Νευρωνικών Δικτύων

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Βρακατσέλης Κωνσταντίνος

Επιβλέπων: Βασίλειος Ασημακόπουλος

Καθηγητής Ε.Μ.Π.



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΩΝ ΔΙΑΤΑΞΕΩΝ & ΣΥΣΤΗΜΑΤΩΝ
ΑΠΟΦΑΣΕΩΝ

Βελτίωση της Μεθόδου Θήτα με χρήση Νευρωνικών Δικτύων

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Επιβλέπων: Βασίλειος Ασημακόπουλος

Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 14/7/2016

.....
Βασίλειος Ασημακόπουλος

Καθηγητής ΕΜΠ

.....
Δημήτριος Ασκούνης

Καθηγητής ΕΜΠ

.....
Ιωάννης Ψαρράς

Καθηγητής ΕΜΠ

Περίληψη

Η εργασία ασχολείται με την εξέταση της αποτελεσματικότητας μιας παραλλαγής της μεθόδου πρόβλεψης Θήτα η οποία κάνει χρήση εργαλείων τεχνητής νοημοσύνης και συγκεκριμένα νευρωνικών δικτύων.

Αρχικά, γίνεται μια ευρεία αλλά συνοπτική επισκόπηση του γενικού πεδίου των προβλέψεων, των εννοιών και των εργαλείων που αυτό χρησιμοποιεί.

Εν συνεχεία, εξετάζεται το θεωρητικό υπόβαθρο της μεθόδου Θήτα στην κλασική της εκδοχή, η πρακτικής εφαρμογής της, καθώς και μια αξιολόγηση της αποτελεσματικότητάς της.

Κατόπιν παρουσιάζονται οι "Διαγωνισμοί Μακριδάκη" ("M-Competitions"), οι οποίοι αποτελούν και το γνωστότερο πεδίο συγκριτικής αξιολόγησης του συνόλου των θεωρητικών και πρακτικών εργαλείων που χρησιμοποιεί το πεδίο των προβλέψεων. Η σκοπιμότητα της παρουσίασης έγκειται στο γεγονός ότι από τους διαγωνισμούς αυτούς αντλείται το πεδίο των χρονοσειρών επί του οποίου θα γίνει τελικά η συγκριτική δοκιμή, ενώ από τους διαγωνισμούς υπάρχει ήδη μια εικόνα σχετικά με τη συγκριτική επίδοση των χρησιμοποιούμενων μεθόδων στην πρόβλεψη.

Προκειμένου να φθάσουμε στην παρουσίαση της παραλλαγής της Μεθόδου Θήτα που χρησιμοποιεί στοιχεία Τεχνητής Νοημοσύνης (μεθόδου "ThetaAI"), γίνεται μια συνοπτική παρουσίαση των νευρωνικών δικτύων, της απαρχής τους καθώς και των τεχνικών που χρησιμοποιούνται για την προσομοίωσή τους. Τα νευρωνικά δίκτυα είναι ο κλάδος του ευρύτερου πεδίου της Τεχνητής Νοημοσύνης που αξιοποιείται για την βελτίωση της επίδοσης της Μεθόδου Θήτα.

Σε λογική συνέχεια, γίνεται η παρουσίαση της παραλλαγής της Μεθόδου Θήτα που εκμεταλλεύεται τεχνικές του πεδίου της Τεχνητής Νοημοσύνης και ειδικότερα των Νευρωνικών Δικτύων. Πρόκειται για τη Μέθοδο ThetaAI (Theta Artificial Intelligence), που είναι μια ήδη επεξεργασμένη και δοκιμασμένη μέθοδος πρόβλεψης σε χρονοσειρές.

Τέλος, γίνεται και πάλι εφαρμογή της μεθόδου ThetaAI σε σύνολα χρονοσειρών που είχαν χρησιμοποιηθεί στους διαγωνισμούς NN3 και M3. Στον διαγωνισμό NN3 η ThetaAI είχε εφαρμοστεί στο σύνολο των 11 χρονοσειρών του διαγωνισμού με μη αυτοματοποιημένη υλοποίηση, όπου και είχε την καλύτερη απόδοση στο διαγωνισμό, και με αυτοματοποιημένο τρόπο στο σύνολο των 11 χρονοσειρών, χωρίς να έχει καλύτερα αποτελέσματα από την κλασική Θήτα. Αυτή τη φορά, εφαρμόζεται η μη αυτοματοποιημένη μέθοδος ThetaAI στις 11 χρονοσειρές του διαγωνισμού NN3 και η αυτοματοποιημένη μέθοδος ThetaAI στις μηνιαίες χρονοσειρές του διαγωνισμού M3.

Abstract

This study examines the effectiveness of a variation of the Theta forecasting method that makes use of artificial intelligence tools, and more specifically of neural networks.

Initially, the study surveys in brief the general field of forecasting and the concept and tools that the field utilizes.

Subsequently, the study examines the theoretical foundation of the Theta method in its classic form, its practical application as well as an appraisal of its effectiveness.

The Makridakis Competitions ("M-Competitions") are then presented, as they constitute the most well-known field of comparative evaluation of the theoretical and practical tools used in the field of forecasting. The rationale of the presentation lies in the fact that the sets of timeseries on which the comparative tests are eventually performed are drawn from them as well as the fact that the M-Competitions already provide a broad view of the comparative effectiveness of forecasting methods.

In order to reach the presentation of the variation of the Theta method that utilizes elements of Artificial Intelligence (the "ThetaAI method") we make a brief presentation of neural networks, their origin as well as the basic techniques used to simulate them. Neural networks constitute a branch of the broader field of Artificial Intelligence and are used to improve the effectiveness of the Theta method.

The variation of the Theta method that exploits artificial intelligence techniques, and more specifically, neural networks, is then presented. It is the Theta AI method, an already established and tested method of forecasting.

Finally, the ThetaAI method is applied again on timeseries sets that had been used in the NN₃ and M₃ competitions. In the NN₃ competition, ThetaAI had been used twice: it was applied once on the 11 timeseries set in a non-automated manner, achieving the best performance in the competition, and once on the 111 timeseries set, without performing better than the classic Theta method. This time, the non-automated ThetaAI method was applied on the 111 timeseries set of the NN₃ competition, and the automated ThetaAI method was applied on the monthly timeseries of the M₃ competition.

Περιεχόμενα

1. Εισαγωγή	9
Δεδομένα και Μέθοδοι Πρόβλεψης	11
Αποσύνθεση.....	13
ARIMA	15
2. Η Μέθοδος Θήτα	16
Το μοντέλο της Μεθόδου Θ	17
Αξιολόγηση της Μεθόδου Θ	21
3. Οι "Διαγωνισμοί Μακριδάκη" (M-Competitions).....	24
4. Νευρωνικά Δίκτυα.....	28
Η Έμπνευση από τη Βιολογία	29
Τροπος Λειτουργίας.....	32
Τα Βήματα Σχεδίασης Νευρωνικού Δικτύου	34
5. Η νευρωνική εκδοχή της Μεθόδου Θήτα	48
Νευρωνικά Δίκτυα και Πρόβλεψη.....	48
Η Μεθοδος Theta AI.	48
6. Βελτίωση των αποτελεσμάτων του διαγωνισμού NN3	52
ΣΥΣΤΗΜΑ ΠΑΡΑΓΩΓΗΣ ΠΡΟΒΛΕΨΕΩΝ ΠΥΘΙΑ	53
ALYUDA NEURO INTELLIGENCE	54
ΒΙΒΛΙΟΘΗΚΗ ALYUDA NEUROFUSION.....	60
7. ΣΥΜΠΕΡΑΣΜΑΤΑ – ΜΕΛΛΟΝΤΙΚΕΣ ΕΠΕΚΤΑΣΕΙΣ	62
Βιβλιογραφία.....	64

1. Εισαγωγή

Στους τομείς της διοίκησης και της διαχείρισης υφίσταται συχνά η ανάγκη για σχεδιασμό, αφού ο χρόνος κατά τον οποίο οι αποφάσεις προηγούνται της δράσης και των αποτελεσμάτων της κυμαίνονται από αρκετά χρόνια (στην περίοδο, πχ, των επενδύσεων κεφαλαίου) έως μερικές ημέρες ή και ώρες (στην περίπτωση σχεδιασμών μεταφορών ή παραγωγής), ακόμη και σε μερικά δευτερόλεπτα (για την δρομολόγηση στις τηλεπικοινωνίες ή την κατανομή ηλεκτρικών φορτίων σε δίκτυα διανομής ενέργειας). Η πρόβλεψη είναι σημαντικό εργαλείο για τον αποδοτικό και αποτελεσματικό σχεδιασμό.

Η ανάγκη για πρόβλεψη αυξάνεται συνεχώς, όσο οι διοικήσεις (management) προσπαθούν όλο και περισσότερο να μειώσουν την εξάρτησή τους από την τύχη με ολοένα και περισσότερο επιστημονικό τρόπο. Καθώς κάθε περιοχή ενδιαφέροντος ενός οργανισμού διασυνδέεται με τις άλλες αντίστοιχες περιοχές, μια καλή ή μία κακή πρόβλεψη μπορεί να επηρεάσει ολόκληρο τον οργανισμό. Μερικές περιοχές ενδιαφέροντος στις οποίες η πρόβλεψη παίζει αυτή τη στιγμή σημαντικό ρόλο είναι:

1. Προγραμματισμός: Η αποδοτική χρήση πόρων απαιτεί τον προγραμματισμό της παραγωγής, των μεταφορών, του προσωπικού και κ.ο.κ. Οι προβλέψεις για το επίπεδο ζήτηση προϊόντων, υλικών, εργασίας, χρηματοδότησης ή υπηρεσιών, είναι ουσιώδη δεδομένα για τον προγραμματισμό.

2. Η απόκτηση συντελεστών παραγωγής: Ο προαπαιτούμενος χρόνος για την απόκτηση πρώτων υλών, προσωπικού ή για την αγορά μηχανημάτων και εξοπλισμού μπορεί να ποικίλει από μερικές ημέρες έως αρκετά χρόνια. Η πρόβλεψη είναι απαραίτητη για να προσδιοριστούν μελλοντικές απαιτήσεις παραγωγικών συντελεστών.

3. Προσδιορισμός των απαιτήσεων συντελεστών παραγωγής: όλοι οι οργανισμοί θα πρέπει να προσδιορίσουν τους πόρους που χρειάζονται μακροπρόθεσμα. Τέτοιου είδους αποφάσεις εξαρτώνται από τις ευκαιρίες των αγορών, περιβαλλοντικούς παράγοντες, καθώς και την ανάπτυξη των ιδίων κεφαλαιακών, ανθρωπίνων, παραγωγικών και τεχνολογικών πόρων. Ο προσδιορισμός αυτός

απαιτεί καλούς υπευθύνους πρόβλεψης και καλούς μάνατζερ που να μπορούν να μεταφράσουν σωστά τις προγνώσεις και να λαμβάνουν τις κατάλληλες αποφάσεις

Η ποσοτική πρόβλεψη μπορεί να εφαρμοστεί όταν ισχύουν τρεις ποσοτικές προϋποθέσεις:

1. Είναι γνωστά δεδομένα από το παρελθόν
2. Τα δεδομένα αυτά μπορούν να ποσοτικοποιηθούν υπό τη μορφή αριθμητικών δεδομένων
3. Μπορεί να γίνει η υπόθεση ότι ορισμένες πτυχές των μορφών και προτύπων συμπεριφοράς του παρελθόντος θα συνεχίσουν να εμφανίζονται και στο μέλλον.

Η τελευταία αυτή προϋπόθεση είναι γνωστή ως υπόθεση συνεχείας. Είναι η υπόρρητη προϋπόθεση για όλες τις ποσοτικές και πολλές ποιοτικές μεθόδους πρόβλεψης, ασχέτως με το πόσο πολύπλοκες μπορεί να είναι.

Οι ποσοτικές μέθοδοι πρόβλεψης ποικίλουν σημαντικά, καθώς έχουν αναπτυχθεί από διαφορετικούς γνωστικούς τομείς και για διαφορετικούς λόγους. Η καθεμία έχει τις δικές της ιδιότητες, ακρίβεια και κόστος που θα πρέπει να συνυπολογιστούν κατά την επιλογή μεθόδου για πρακτική χρήση. Οι ποσοτικές διαδικασίες πρόβλεψης βρίσκονται σε ένα συνεχές που ορίζεται από δύο άκρα: τις διαισθητικές ή ad hoc μεθόδους, και τις τυπικές ποσοτικές μεθόδους που βασίζονται στις αρχές της στατιστικής. Ο πρώτος τύπος βασίζεται στην πρακτική εμπειρία που ποικίλει από εταιρεία σε εταιρεία, προϊόν σε προϊόν και υπεύθυνο προβλέψεων σε υπεύθυνο προβλέψεων. Οι διαισθητικές μέθοδοι είναι απλές και εύχρηστες, αλλά συχνά όχι τόσο ακριβείς όσο οι τυπικές ποσοτικές μέθοδοι. Επίσης, συχνά δεν παρέχουν καμία πληροφορία σχετικά με το περιθώριο ακριβείας της πρόβλεψης. Εξ αιτίας των περιορισμών αυτών, η χρήση τους φθίνει, ενώ έχουν αντιστρόφως διαδοθεί οι τυπικές μέθοδοι πρόβλεψης. Πολλές επιχειρήσεις χρησιμοποιούν ακόμη τις μεθόδους αυτές είτε επειδή δεν γνωρίζουν σχετικά με τις απλές τυπικές μεθόδους είτε επειδή προτιμούν μια προσέγγιση που βασίζεται στην κρίση παρά στα αντικειμενικά δεδομένα.

Δεδομένα και Μέθοδοι Πρόβλεψης

Κλασική Μεθοδολογία Πρόβλεψης

Η κλασική μεθοδολογία της Πρόβλεψης περιλαμβάνει ως εργαλεία δύο βασικές και διακριτές υποδιαίρεσεις:

Αιτιοκρατικό Μοντέλο

Στο αιτιοκρατικό μοντέλο της Πρόβλεψης θεωρείται ότι το μέγεθος του οποίου ενδιαφέρει η πρόβλεψη της εξέλιξης, εξαρτάται κατά καθορισμένο, "αιτιοκρατικό" τρόπο από άλλες μεταβλητές. Με άλλα λόγια, ο τρόπος που διαμορφώνεται ανά πάσα στιγμή η τιμή ενός μεγέθους, εξαρτάται από τις τιμές που έχουν κατά την ίδια στιγμή, οι μεταβλητές που – αιτιοκρατικώς – καθορίζουν την τιμή του μεγέθους που ενδιαφέρει. Έτσι, για να προβλέψει κανείς τη μελλοντική τιμή του μεγέθους ενδιαφέροντος, αρκεί να γνωρίζει τις μελλοντικές τιμές των παραγόντων που το προσδιορίζουν. Με δεδομένες τις τιμές των μεταβλητών αυτών, καθώς και τον "μηχανισμό διαμόρφωσης" του μεγέθους (μαθηματικό ή άλλου είδους), μπορεί κανείς να υπολογίσει τις μελλοντικές τιμές που ενδιαφέρουν.

Τα βασικά προβλήματα με τη χρήση των αιτιοκρατικών μεθόδων είναι δύο:

- ✓ Αφ' ενός, για πολλά φαινόμενα που ενδιαφέρουν, δεν είναι δυνατή η δημιουργία του αιτιοκρατικού "μοντέλου" δημιουργίας τους. Αν και αιτιοκρατικά και όχι τυχαία, ο μηχανισμός διαμόρφωσης είναι μαθηματικά τόσο πολύπλοκος που είτε μπορεί να θεωρηθεί άγνωστος είτε πρακτικά δύσχρηστος.
- ✓ Αφ' ετέρου, οι μελλοντικές τιμές των ιδίων των μεταβλητών που διαμορφώνουν αιτιοκρατικά το μέγεθος που μας ενδιαφέρει, είναι συχνά άγνωστες. Έτσι, ακόμη κι αν ο μηχανισμός είναι γνωστός, παραμένει χωρίς χρησιμότητα ως προς τη δυνατότητα μελλοντικών προβλέψεων.

Ανάλυση Χρονοσειρών

Στην περίπτωση αυτή της Ανάλυσης Χρονοσειρών, η προσπάθεια για την πρόβλεψη των μελλοντικών τιμών ενός μεγέθους δε λαμβάνει καθόλου τους εξωτερικούς παράγοντες-ανεξάρτητες μεταβλητές που επηρεάζουν το μέγεθος αυτό. Αντιθέτως, το σκεπτικό της προσέγγισης αυτής έγκειται στην υπόθεση ότι, ανεξαρτήτως από τους παράγοντες που ανά πάσα στιγμή διαμορφώνουν αιτιοκρατικά ένα μέγεθος, το ίδιο το μέγεθος ακολουθεί μια "εσωτερική κανονικότητα" στη χρονική του εξέλιξη. Συνεπώς, η μέθοδος αυτή επιχειρεί να εντοπίσει την "εσωτερική κανονικότητα" του φαινομένου βάσει των παρελθουσών τιμών του μεγέθους. Η μέθοδος στηρίζεται δηλαδή στην υπόθεση ότι από την εξέλιξη των προηγούμενων τιμών του μεγέθους μπορούν να εξαχθούν οι επόμενες, χωρίς προσφυγή σε άλλα μεγέθη που εξωγενώς καθορίζουν το υπό εξέταση μέγεθος.

Από τεχνικής πλευράς, η μέθοδος βασίζεται στα στατιστικά δεδομένα που παρέχουν οι παρελθούσες τιμές του προς εκτίμηση μεγέθους. Όμως, η μέθοδος της Ανάλυσης Χρονοσειρών (Time Series Analysis) δεν αποτελεί απλή εφαρμογή της Παλινδρόμησης, που είναι το προφανές εργαλείο, δηλαδή δεν αποτελεί απλή - γραμμική ή μη - προβολή των προηγούμενων τιμών στις επόμενες.

Αντιθέτως, η Ανάλυση Χρονοσειρών έγκειται στην ανίχνευση και προσδιορισμό των εσωτερικών μοντέλων (patterns) που ακολουθούν οι τιμές του μεγέθους, προκειμένου αυτές να δώσουν μελλοντικές τιμές. Για την ακρίβεια, η Ανάλυση Χρονοσειρών **υποθέτει** την ύπαρξη βασικών εσωτερικών μοντέλων (patterns) στην εξέλιξη ενός μεγέθους, και εν συνεχεία επιδιώκει να τα προσδιορίσει ποσοτικά. Η υπόθεση αυτή είναι αυστηρά περιγραφική και όχι αιτιοκρατική: ο μηχανισμός είναι ένα "μαύρο κουτί" που μας επιτρέπει να προβλέπουμε χωρίς να μας επιτρέπει να κατανοούμε το γιατί συμβαίνει ό,τι συμβαίνει.

Η προσέγγιση δίνει τα τρία βασικότερα εργαλεία-μεθόδους ανάλυσης χρονοσειρών:

- **Μέθοδο Αποσύνθεσης**
- **Μέθοδος Εξομάλυνσης**
- **Μέθοδο ARIMA**

Φυσικά, αυτές αποτελούν απλώς γενικές οικογένειες μεθόδων, που στην πράξη εξειδικεύονται και βελτιώνονται με πληθώρα παραλλαγών.

Αποσύνθεση

Στην γενική μέθοδο της αποσύνθεσης γίνεται η υπόθεση ότι η χρονοσειρά αναλύεται σε τέσσερις βασικές συνιστώσες:

- **Τάση**
- **Κύκλος**
- **Εποχιακότητα**
- **Τυχειότητα**

Η μέθοδος της αποσύνθεσης έγκειται στον προσδιορισμό των τεσσάρων αυτών συνιστωσών με βάση τις παρελθούσες τιμές, και την εφαρμογή τους προκειμένου να γίνει πρόβλεψη των επομένων τιμών.

Σε ότι αφορά τις τέσσερις αυτές συνιστώσες:

Η **Τάση** (*Trend* – T_t): Αποτελεί τη συστηματική μεταβολή του μέσου επιπέδου της χρονοσειράς ως συνάρτηση του χρόνου.

Η **Εποχικότητα** (*Seasonality* – S_t): Αποτελεί τη συστηματική περιοδική διακύμανση (δηλαδή σταθερού χρονικού διαστήματος) που γενικά συμπίπτει με τις περιόδους του έτους ή πάντως με σταθερά, δεδομένα χρονικά διαστήματα.

Η **Κυκλικότητα** (*Cycle* – C_t): Αποτελεί επίσης μια περιοδικότητα, μεγαλύτερης περιόδου από την εποχικότητα, ενώ η περίοδος της δεν έχει εκ των προτέρων γνωστή ή αναμενόμενη διάρκεια, ή σχετιζόμενη με κάποια δεδομένη, εξωτερική περίοδο.

Η **Τυχειότητα** (*Randomness* – R_t): Αποτελεί τη συνιστώσα που απομένει μετά την ταυτοποίηση των υπολοίπων τριών βασικών συνιστωσών και εκφράζει τη μη-προβλέψιμη και μη-συστηματική μεταβολή του επιπέδου τιμών της χρονοσειράς.

Από μαθηματικής απόψεως, θεωρούμε ότι η χρονοσειρά είναι συνάρτηση των τεσσάρων αυτών μεταβλητών:

$$X_t = f(T_t, S_t, C_t, R_t)$$

Γενικώς, για την ανάλυση της χρονοσειράς γίνεται υπόθεση δύο δυνατών μορφών της συνάρτησης:

A. Προσθετική Μορφή

Στην περίπτωση αυτή, θεωρείται ότι η χρονοσειρά αποτελεί συνάρτηση της μορφής:

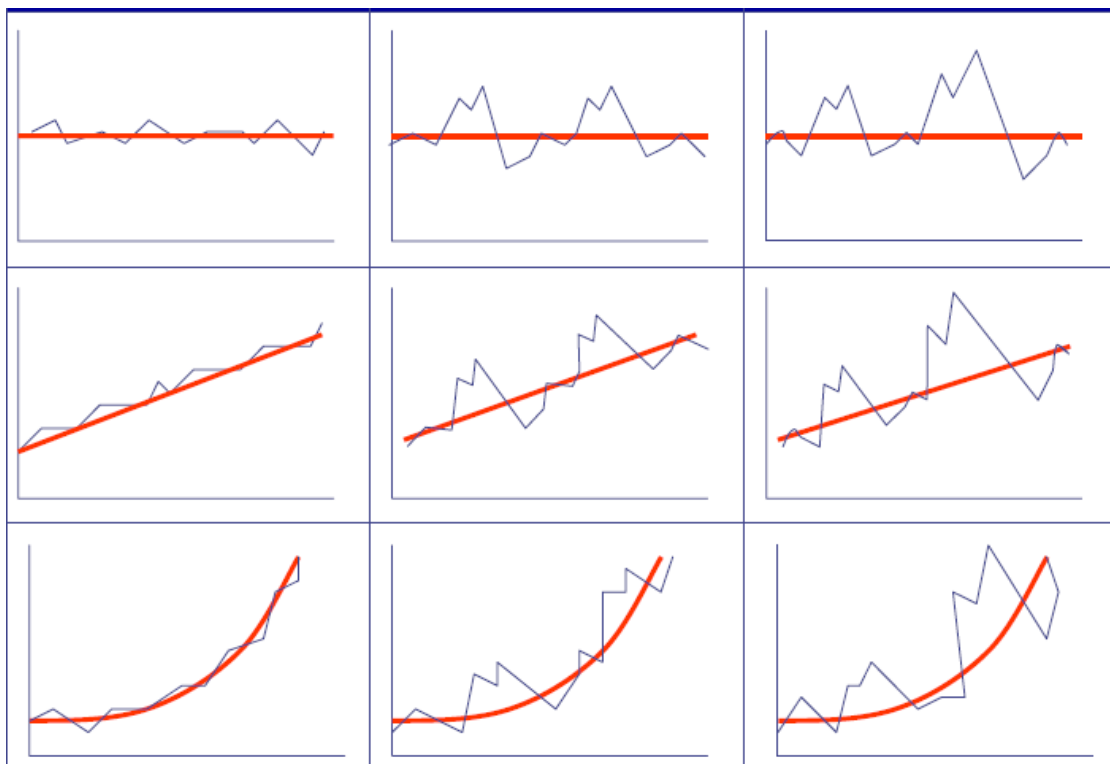
$$X_t = T_t + C_t + R_t + S_t$$

B. Πολλαπλασιαστική Μορφή

Στην περίπτωση αυτή, θεωρείται ότι η χρονοσειρά αποτελεί συνάρτηση της μορφής:

$$X_t = T_t * C_t * R_t * S_t$$

Ο συνδυασμός των συνιστωσών στα δύο μοντέλα φαίνεται χαρακτηριστικά στο σχήμα:



ARIMA

Η μέθοδος ARIMA (Auto Regressive Integrated Moving Average – Αυτοπαλινδρομική Μέθοδος Ολοκληρωμένων Κινητών Μέσων Όρων) αποτελεί μία στοχαστική μέθοδο πρόβλεψης φαινομένων.

Εννοιολογικά, η υπεροχή της έναντι της απλής αποσύνθεσης έγκειται στο ότι, σε αντίθεση με την απλή αποσύνθεση που δίνει την ίδια βαρύτητα στο σύνολο των διαθέσιμων τιμών της χρονοσειράς, όσο πρόσφατες ή παλιές κι αν είναι, η μέθοδος ARIMA δίνει μεγαλύτερη βαρύτητα στις πρόσφατες τιμές έναντι των παλαιότερων, με αποτέλεσμα μεγαλύτερη αξιοπιστία. Από την άλλη, ακριβώς εξ αιτίας της ιδιαίτερης βαρύτητας των πρόσφατων τιμών, η μέθοδος ARIMA και οι παραλλαγές της παρουσιάζουν καλύτερη επίδοση στις βραχυπρόθεσμες προβλέψεις ενώ η επίδοση είναι λιγότερο ικανοποιητική στον μακροπρόθεσμο ορίζοντα.

Από τεχνικής απόψεως, τα βήματα της μεθόδου ARIMA έχουν προσδιοριστεί από τους Box και Jenkins, στο πλαίσιο της γενικότερης μοντελοποίησης μεθόδου χρονοσειρών.

Συγκεκριμένα, η μέθοδος ακολουθεί διαδικαστικά τρία στάδια:

Ταυτοποίηση (identification) του μοντέλου που θα χρησιμοποιηθεί: Δεδομένου ότι η χρονοσειρά θα πρέπει να είναι στάσιμη (stationary), η δοθείσα αρχική χρονοσειρά μετασχηματίζεται με την εφαρμογή διαδοχικών διαφορών (με άλλα λόγια, "αφαιρείται" η τάση από το μοντέλο, με χρήση συναρτήσεων αυτοσυσχέτισης προκειμένου να προσδιοριστεί η έκταση στην οποία πρέπει να εφαρμοστούν διαδοχικές διαφορές.

Υπολογισμός (estimation) και **έλεγχος** (testing) του μοντέλου που έχει επιλεγεί:

Εφαρμογή (application) του προσδιορισμένου μοντέλου: Αφού το μοντέλο έχει προσδιοριστεί, εισάγονται πάλι τα δεδομένα της χρονοσειράς, προκειμένου να εξαχθούν οι τα δεδομένα της πρόβλεψης.

2. Η Μέθοδος Θήτα

Μία διαφορετική μέθοδος πρόβλεψης, και συγκεκριμένα μέθοδος αποσύνθεσης με την επωνυμία "Μέθοδος Θήτα" παρουσιάστηκε το 2000 από τους Ασημακόπουλο και Νικολόπουλο.

Στην κλασική μέθοδο της αποσύνθεσης, η χρονοσειρά αναλύεται σε συνιστώσες τάσης-κύκλου, εποχιακότητας και τυχειότητας. Οι συνιστώσες προεκτείνονται στο μέλλον και επανασυντίθενται, δίνοντας το σύνολο των μελλοντικών τιμών της σειράς. Σε αντιδιαστολή με αυτή την προσέγγιση, η μέθοδος θήτα, αφού αφαιρέσει την εποχικότητα από τη χρονοσειρά, αποσυνθέτει τα δεδομένα σε δύο συνιστώσες: σε μία βραχυπρόθεσμη και μία μακροπρόθεσμη τάση. Η κάθε μία από αυτές τις δύο τάσεις

Το μοντέλο της Μεθόδου Θ

Η μέθοδος θήτα βασίζεται στη μεταβολή των τοπικών καμπυλοτήτων της χρονοσειράς επί της οποίας εφαρμόζεται. Αυτό επιτυγχάνεται με τη χρήση του "συντελεστή θ". Ο συντελεστής αυτός εφαρμόζεται στις δεύτερες διαφορές των δεδομένων της χρονοσειράς, σύμφωνα με τη σχέση

$$X''_{new}(\theta) = \theta \cdot X''_{data}$$

όπου

$$X''_{data} = X_t - 2X_{t-1} + X_{t-2}, \text{ σε χρόνο } t$$

Αν σταδιακά μειώνονται οι τοπικές καμπυλότητες, τότε η χρονοσειρά εξομαλύνεται, όπως φαίνεται στο ακόλουθο σχήμα:

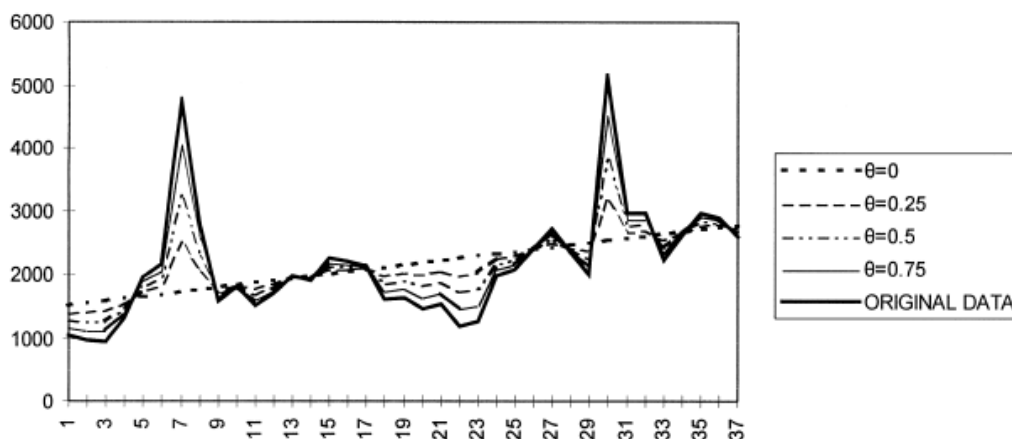


Fig. 1. M3-Comp. Series 200, the Theta-model deflation.

Όσο μικρότερη είναι η τιμή του συντελεστή θήτα, τόσο μεγαλύτερη είναι η εξομάλυνση, δηλαδή τόσο περισσότερο μειώνονται οι τοπικές καμπυλότητες. Για την τιμή $\theta=0$, η χρονοσειρά μετατρέπεται στην ευθεία γραμμή της γραμμικής παλινδρόμησης. Η προοδευτική μείωση στις διακυμάνσεις οδηγεί στην ελάττωση των διαφορών μεταξύ των διαφορετικών παρατηρήσεων. Ποιοτικά, το αποτέλεσμα αυτό συνδέεται με την ανάδειξη των μακροπρόθεσμων τάσεων στα δεδομένα.

Είναι δυνατόν ο συντελεστής θ να λάβει και αρνητικές τιμές, οι οποίες οδηγούν στην αντιστροφή της τοπικής καμπυλότητας στην χρονοσειρά επί της οποίας εφαρμόζεται.

Εάν, αντιστρόφως, η τοπική καμπυλότητα αυξηθεί, δηλαδή για $\theta > 1$, τότε η χρονοσειρά "εκτείνεται", όπως φαίνεται στο ακόλουθο σχήμα:

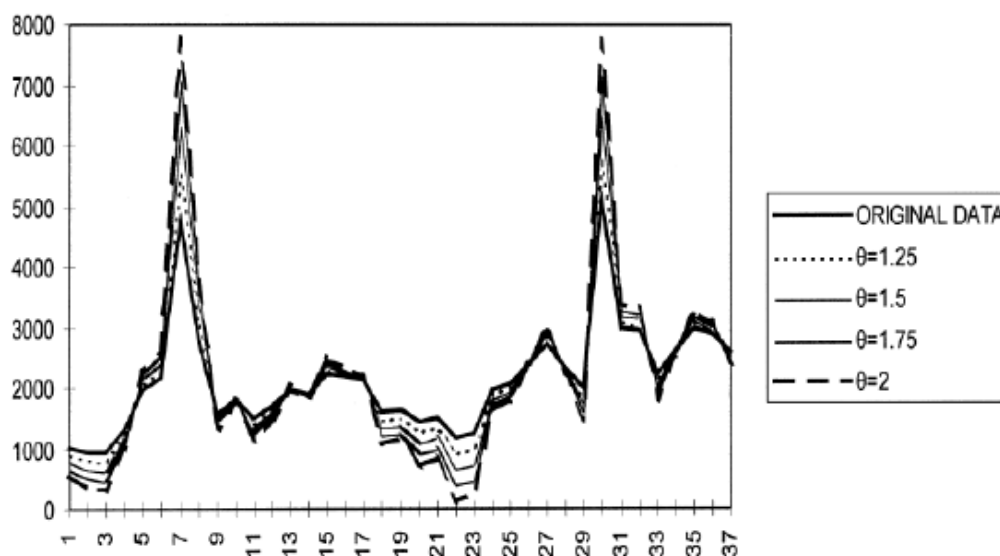


Fig. 2. M3-Comp. Series 200, the Theta-model dilation.

Από το σχήμα γίνεται διαισθητικά αντιληπτό πως ο μεγαλύτερος βαθμός "έκτασης", (δηλαδή η μεγαλύτερη τιμή του θ), οδηγεί σε "μεγέθυνση" της βραχυπρόθεσμης συμπεριφοράς της χρονοσειράς. Για τη χαρακτηριστική τιμή $\theta = -1$,

Με βάση την ανωτέρω διαδικασία, είναι δυνατόν να δημιουργηθεί ένα σύνολο από γραμμές, οι οποίες ονομάζονται "γραμμές Θήτα". Η τοποθέτηση των γραμμών αυτών σε σχέση με τα δεδομένα της χρονοσειράς μπορούν να γίνουν με διάφορους τρόπους. Σε περίπτωση που οι γραμμές διαμορφωθούν σύμφωνα με τη μέθοδο των ελαχίστων τετραγώνων, διατηρείται η μέση τιμή και η κλίση της αρχικής χρονοσειράς.

Η γενική σχηματοποίηση της μεθόδου έχει ως εξής:

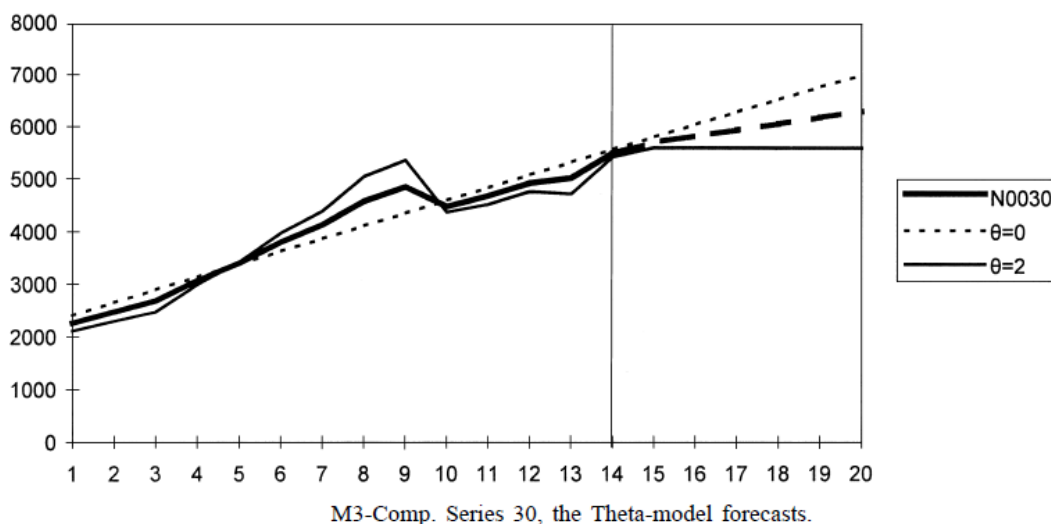
Η αρχική χρονοσειρά αποσυντίθεται σε δύο ή περισσότερες γραμμές θ . Αυτές προεκτείνονται στο μέλλον ξεχωριστά και απλά συνδυάζονται οι προβλέψεις με ίσα βάρη. Μπορεί να χρησιμοποιηθεί οποιαδήποτε μέθοδος πρόβλεψης για να παραχθούν προβλέψεις για την κάθε γραμμή Θ , με βασικό κριτήριο την εμπειρία που έχει προκύψει για την συγκεκριμένη χρονοσειρά. Ένας διαφορετικός συνδυασμός των γραμμών Θ μπορεί να εφαρμοστεί για κάθε ορίζοντα πρόβλεψης.

Ας θεωρήσουμε μια από τις πιο απλές περιπτώσεις όπου η αρχική χρονοσειρά αποσυντίθεται σε δύο γραμμές Θ , για τιμές του συντελεστή θ , $\theta=0$ και $\theta=2$:

$$\text{Data} = \frac{1}{2} (L(\theta=0) + (L(\theta=2)))$$

όπου $L(\theta=a)$ είναι η γραμμή Θ για τιμή του συντελεστή θ ίση με a .

Η πρώτη γραμμή Θ ($\theta=0$) είναι η ευθεία της απλής γραμμικής παλινδρόμησης των δεδομένων και η δεύτερη έχει δεύτερες διαφορές διπλάσιες της αρχικής χρονοσειράς. Αυτή είναι η περίπτωση όπου δύο γραμμές Θ , με τιμές παραμέτρου θ συμμετρικές ως προς τη μονάδα, συνθέτονται. Η πρώτη γραμμή ($L(\theta=0)$) περιγράφει την χρονοσειρά μέσα από μια γραμμική τάση. Η δεύτερη ($L(\theta=2)$) έχει διπλασιάσει τις τοπικές καμπυλότητες τονίζοντας έτσι την βραχυπρόθεσμη συμπεριφορά. Η πρώτη γραμμή προεκτείνεται στο μέλλον σύμφωνα με την κλασική θεωρία παλινδρόμησης. Αντίστοιχα η δεύτερη προεκτείνεται με εκθετική εξομάλυνση σταθερού επιπέδου SES. Ο απλός γραμμικός συνδυασμός των δύο προβλέψεων δίνει την τελική πρόβλεψη για το μοντέλο Θ . Αυτή η διαδικασία απεικονίζεται στο παρακάτω σχήμα:



Τα βήματα που ακολουθούνται κατά την εφαρμογή της μεθόδου αναλυτικά είναι τα εξής:

ΒΗΜΑ 0: Τέστ Εποχιακότητας

Η κάθε χρονοσειρά ελέγχεται για στατιστικώς σημαντική εποχιακή συμπεριφορά. Κριτήριο είναι η τιμή της t-κατανομής με πιθανότητα 0.1 τιμή για την τιμή συνάρτησης αυτοσυσχέτισης με καθυστέρηση ένα έτος (δηλαδή για 12 παρατηρήσεις αν πρόκειται για μηνιαία δεδομένα ή 4 παρατηρήσεις αν πρόκειται για τριμηνιαία). Η σύγκριση γίνεται με την τιμή 1.645, που είναι η τιμή για πιθανότητα 0,1.

ΒΗΜΑ 1: Αποεποχικοποίηση

Η χρονοσειρά αποεποχικοποιείται με την κλασσική μέθοδο αποσύνθεσης

ΒΗΜΑ 2: Αποσύνθεση

Η χρονοσειρά αποσυντίθεται σε δύο γραμμές Θ , για τιμές $\theta=0$ και $\theta=2$.

ΒΗΜΑ 3: Πρόβλεψη

Η γραμμή $L(\theta=0)$ προεκτείνεται με απλή γραμμική παλινδρόμηση ενώ η $L(\theta=2)$ με εκθετική εξομάλυνση σταθερού επιπέδου.

ΒΗΜΑ 4: Συνδυασμός

Οι προηγούμενες προβλέψεις συνδυάζονται με ίσα βάρη.

ΒΗΜΑ 5: Εποχικοποίηση

Οι τελικές προβλέψεις εποχικοποιούνται.

Αξιολόγηση της Μεθόδου Θ

Το δυνατό στοιχείο της μεθόδου έγκειται στην αποσύνθεση των αρχικών δεδομένων. Οι δύο συνιστώσες περιέχουν πληροφορίες, οι οποίες είναι χρήσιμες για την διαδικασία πρόβλεψης αλλά χάνονται ή δεν μπορούν να αναγνωριστούν και να αξιοποιηθούν από τα υπάρχοντα μοντέλα πρόβλεψης όταν αυτά εφαρμόζονται απευθείας στα αρχικά δεδομένα. Ιδιαίτερα στην περίπτωση της γραμμής $L(\theta=0)$ το φαινόμενο αυτό είναι πιο εύκολα αντιληπτό. Η ευθεία γραμμή παλινδρόμησης περιέχει πληροφορίες για την μακροπρόθεσμη τάση της χρονοσειράς η οποία αμελείται όταν η μέθοδος προσαρμόζεται σε πιο πρόσφατες τάσεις. Από την άλλη πλευρά, όταν χρησιμοποιείται μόνο η γραμμική τάση, όλες οι πολύτιμες πληροφορίες των βραχυχρόνιων διακυμάνσεων αγνοούνται.

Χαρακτηριστικό παράδειγμα είναι η επίδοση του μοντέλου Θ στις μηνιαίες χρονοσειρές του διαγωνισμού M3. Τα μηνιαία δεδομένα του διαγωνισμού έχουν χαρακτηριστεί γενικά από σχετικά μεγάλες διακυμάνσεις. Αυτό το γεγονός δεν επιτρέπει στις περισσότερες μεθόδους να αναγνωρίσουν την μακροπρόθεσμη τάση και κατά συνέπεια να την λάβουν σοβαρά υπόψη τους στην παραγωγή προβλέψεων. Στη περίπτωση του μοντέλου Θ, η μακροπρόθεσμη τάση εξασφαλίζεται από την προέκταση της γραμμής $\theta=0$. Ταυτόχρονα η ύπαρξη και της γραμμής $\theta=2$ λειτουργεί σαν αντίβαρο στην χρησιμοποίηση μόνο της απλής γραμμικής παλινδρόμησης και εξασφαλίζει την αξιοποίηση και της βραχυπρόθεσμης πληροφορίας. Ως αποτέλεσμα το σημείο εκκίνησης των προβλέψεων πετυχαίνει καλύτερη προσέγγιση του σωστού επιπέδου, και μια και η προέκταση είναι οριζόντια εξασφαλίζεται μια συντηρητική μεν σταθερή δε συνέχιση της μακροπρόθεσμης τάσης.

Αναλυτική μαθηματική παρουσίαση της μεθόδου θ βρίσκεται στα παραρτήματα του αρχικού άρθρου με το οποίο παρουσιάστηκε στην επιστημονική κοινότητα η μέθοδος, που δημοσιεύτηκε στο *International Journal of Forecasting* (2000):

*“The Theta Model: a decomposition approach to forecasting” V.Assimakopoulos ,
K.Nikolopoulos.*

3. Οι "Διαγωνισμοί Μακρινάκη" (M-Competitions)

Για τον ρεαλιστικό και πρακτικό έλεγχο της επιτυχίας των μεθόδων πρόβλεψης, από το 1982 καθιερώθηκαν και διοργανώνονται διεθνείς επιστημονικοί διαγωνισμοί, με την επωνυμία "Διαγωνισμοί Μακρινάκη" ή Διαγωνισμοί "M" (M-Competitions).

Οι διαγωνισμοί καθιερώθηκαν και οργανώνονται υπό την εποπτεία του ειδικού σε θέματα προβλέψεων καθηγητή Σπύρου Μακρινάκη. Αφορμή υπήρξε άρθρο των Μακρινάκη και Hibon ("*Accuracy of Forecasting: An Empirical Investigation*") με το οποίο οι συντάκτες υποστήριζαν ότι οι μαθηματικά πολύπλοκες στατιστικές μέθοδοι πρόβλεψης δεν αποδίδουν καλύτερα αποτελέσματα έναντι των μαθηματικά απλούστερων αντίστοιχων μεθόδων. Το άρθρο προκάλεσε αντιδράσεις στην κοινότητα των στατιστικολόγων, και ο Μακρινάκης διοργάνωσε τον πρώτο διαγωνισμό ώστε να ελεγχθεί στην πράξη ένα μεγάλο πλήθος μεθόδων πρόβλεψης με βάση τα ίδια δεδομένα (χρονοσειρές) προκειμένου να συγκριθούν αξιόπιστα τα αποτελέσματά τους, στον διαγωνισμό που έχει αποκληθεί M-1.

Σύμφωνα με τους διοργανωτές, ο πρώτος διαγωνισμός, που χρησιμοποίησε 1001 χρονοσειρές και ήλεγξε 15 μεθόδους πρόβλεψης, μαζί με εννέα ακόμη παραλλαγές των μεθόδων αυτών, οδήγησε στα εξής τρία συμπεράσματα:

- Οι στατιστικώς εξεζητημένες ή σύνθετες μέθοδοι δεν παρέχουν κατ' ανάγκη πιο ακριβείς προβλέψεις από τις απλούστερες
- Η σχετική κατάταξη της επίδοσης των διαφόρων μεθόδων ποικίλει ανάλογα με το μέτρο της ακρίβειας που χρησιμοποιείται
- Σε γενικές γραμμές, η χρήση συνδυασμού διαφορετικών μεθόδων αποδίδει καλύτερα αποτελέσματα από συνδυασμένες μοναδικές μεθόδους και αποδίδει πολύ καλά συγκριτικά με άλλες μεθόδους
- Η ακρίβεια των διαφόρων μεθόδων εξαρτάται από το μήκος του ορίζοντα πρόβλεψης που εμπλέκεται.

Ο διαγωνισμός επαναλήφθηκε το 1993, με την επωνυμία M-2, και με τροποποιημένη μεθοδολογία. Συγκεκριμένα, ο διαγωνισμός ήταν σημαντικά

μεγαλύτερης κλίμακας ως προς τους συμμετέχοντες, αλλά πιο περιορισμένος και εστιασμένος ως προς τον αριθμό των χρονοσειρών που ελέγχθηκαν, οι οποίες επελέγησαν με αυστηρότερα κριτήρια. Ειδικότερα, έναντι των 1001 χρονοσειρών που ελέγχθηκαν στον M-1, στον δεύτερο διαγωνισμό ελέγχθηκαν 29 συνολικά χρονοσειρές, εκ των οποίων οι 23 προέρχονταν από συνεργαζόμενες με τη διοργάνωση του διαγωνισμού εταιρείες, ενώ 6 χρονοσειρές περιείχαν μακροοικονομικά δεδομένα. Ο σκοπός της τροποποίησης της μεθοδολογίας στο διαγωνισμό M-2 ήταν να γίνει πιο ρεαλιστική προσομοίωση των διαδικασιών πρόβλεψης όπως αυτές εφαρμόζονται στην πράξη, και για τον λόγο αυτόν:

- Επιτρεπόταν στους υπευθύνους προβλέψεων να συνδυάσουν τη μέθοδο πρόβλεψης που βασιζόταν στην τάση με την προσωπική τους κρίση
- Επιτρεπόταν στους υπευθύνους προβλέψεων να θέσουν πρόσθετα διευκρινιστικά ερωτήματα στους παρόχους των χρονοσειρών (δηλαδή στις εταιρείες) προκειμένου να βελτιώσουν τις προβλέψεις τους

Επιτρεπόταν στους υπευθύνους προβλέψεων να ενσωματώνουν τα διδάγματα από κάθε μεμονωμένη πρόβλεψη στην εφαρμογή της προβλέψεώς τους για την επόμενη χρονοσειρά, βασιζόμενοι στην ανατροφοδότηση.

Ο διαγωνισμός έτυχε ευρείας αποδοχής και τα αποτελέσματά του δημοσιεύτηκαν σε εργασία, το 1993. Τα αποτελέσματα του διαγωνισμού M-2 συνέπιπτα στατιστικά με τα αποτελέσματα του διαγωνισμού M-1, επιβεβαιώνοντας και τα αρχικά συμπεράσματα. Παρ' όλα αυτά, οι Μακρινδάκης και Fildes υποστηρίζουν ότι στην τα συμπεράσματα των δύο διαγωνισμών τα αγνοεί στην πρακτική της.

Το 2000 διεξήχθη ο τρίτος διαγωνισμός της σειράς, με την επωνυμία M-3. Σκοπός του τρίτου διαγωνισμού ήταν όχι απλώς να επιβεβαιώσει τα αποτελέσματα των προηγούμενων δύο αλλά και να διευρύνει τη σχετική εμπειρία. Για τον λόγο αυτό, ο αριθμός των ελεγχόμενων χρονοσειρών ανέβηκε και πάλι, και ανήλθε στις 3003. Επιπλέον, αυξήθηκε σημαντικά τόσο ο αριθμός των ερευνητών που έλαβαν μέρος όσο και ο αριθμός των μεθόδων πρόβλεψης που χρησιμοποιήθηκαν. Τέλος, σε μια

προσπάθεια ουσιώδους εμπλουτισμού του διαγωνισμού, έγινε πρόβλεψη (και πρόσκληση) για συμμετοχή σε μεθόδους βασιζόμενες σε Νευρωνικά Δίκτυα.

Οι χρονοσειρές που δόθηκαν στον διαγωνισμό αποτελούσαν ένα ευρύ μείγμα από ετήσιες, τριμηνιαίες, μηνιαίες, ημερήσιες και άλλου είδους χρονοσειρές. Προκειμένου να επιτευχθεί ένα ελάχιστο επίπεδο ακρίβειας του μοντέλου πρόβλεψης, οι χρησιμοποιούμενες χρονοσειρές είχαν ένα ελάχιστο παρατηρήσεων: 14 για τις ετήσιες, 16 για τις τριμηνιαίες, 48 για τις μηνιαίες, και 60 παρατηρήσεις για οποιαδήποτε άλλης χρονοσειρά. Η προέλευσή τους ήταν επίσης ποικίλη: υπήρχαν χρονοσειρές μικροοικονομικής, μακροοικονομικής, βιομηχανικής, δημογραφικής και άλλων προελεύσεων.

Για τον έλεγχο και τη σύγκριση των αποτελεσμάτων χρησιμοποιήθηκαν πέντε διαφορετικά μέτρα:

- Το (συμμετρικό) μέσο απόλυτο ποσοστιαίο σφάλμα (*mean average percentage error – MAPE*),
- Μέση κατάταξη (*average ranking*)
- Βελτίωση ποσοστού – έναντι άλλων μεθόδων (*percentage better*)
- Διάμεσο συμμετρικό απόλυτο ποσοστιαίο σφάλμα (*median symmetric Absolute Percentage Error*)
- Μέσο σχετικό απόλυτο σφάλμα (*median Relative Absolute Error*)

Η επεξεργασία των αποτελεσμάτων του διαγωνισμού επιβεβαίωσε τα βασικά συμπεράσματα των δύο προηγούμενων διαγωνισμών. Επισημαίνεται ότι στον διαγωνισμό αυτόν έλαβε μέρος για πρώτη φορά και η μέθοδος Θήτα των Ασημάκη-Νικολόπουλου, όπου και είχε εξαιρετική επίδοση.

Χωρίς να ανήκει στην σειρά των Διαγωνισμών Μακριδάκη, αλλά άμεσο απότοκό τους υπήρξε ο Διαγωνισμός NN-3 που οργάνωσαν το 2005 οι Crone, Νικολόπουλος και Hibon. Σκοπός του διαγωνισμού ήταν να προκαλέσει μεγαλύτερη συμμετοχή νευρωνικών μεθόδων και να εξαγάγει συμπερασμάτων σχετικά με την προγνωστική τους ισχύ. Αν και ο διαγωνισμός M-3 είχε επίσης επιδιώξει τη συμμετοχή νευρωνικών μεθόδων πρόβλεψης, η συμμετοχή ήταν πενιχρή (για την ακρίβεια μία και μοναδική) – πιθανότατα εξ αιτίας των μεγάλων

υπολογιστικών απαιτήσεων των μεθόδων αυτών, σε συνδυασμό με τον μεγάλο όγκο των δεδομένων των παρεχόμενων γραμματοσειρών. Προκειμένου στον διαγωνισμό αυτόν να υπάρχει ομοιογένεια με τα δεδομένα του M-3 ώστε η σύγκριση να είναι άμεση και ευχερής, χρησιμοποιήθηκαν 111 από τις χρονοσειρές του M-3, με την χρονική μετάθεση των δεδομένων που εν τω μεταξύ είχαν προκύψει.

- Ο διαγωνισμός NN-3 είχε επαρκή συμμετοχή ώστε να επιτρέπει την εξαγωγή βασικών συμπερασμάτων σχετικά με τις νευρωνικές μεθόδους πρόβλεψης. Τα κεντρικά συμπεράσματα ήταν
- Οι καλύτερες νευρωνικές μέθοδοι είχαν αποτελεσματικότητα συγκρίσιμη με αυτή των καλύτερων κλασικών μεθόδων πρόβλεψης
- Οι νευρωνικές απαιτήσεις είχαν πολύ μεγαλύτερες υπολογιστικές απαιτήσεις από τις αντίστοιχες κλασικές
- Πολλές νευρωνικές μέθοδοι στην πράξη απέδιδαν πολύ λιγότερο από το θεωρητικώς προβλεπόμενο, και υπολείπονταν σημαντικά ακόμη και κλασικών μεθόδων πρόβλεψης, παρά τις θεωρητικά υψηλές προσδοκίες απόδοσης.

4. Νευρωνικά Δίκτυα

Σύμφωνα με τον κλασικό ορισμό του Haykin, Νευρωνικά Δίκτυα είναι:

Μεγάλος παράλληλος κατανεμημένος επεξεργαστής αποτελούμενος από απλές μονάδες επεξεργασίας που έχει τη φυσική ιδιότητα να αποθηκεύει εμπειρική γνώση και να την καθιστά διαθέσιμη για χρήση. Προσομοιάζει με τον ανθρώπινο εγκέφαλο από δύο απόψεις:

- 1. Το νευρωνικό δίκτυο αποκτά τη γνώση από το περιβάλλον μέσω μιας διαδικασίας εκμάθησης*
- 2. Η ισχύς των συνδέσεων μεταξύ των απλών μονάδων επεξεργασίας ("νευρώνων"), γνωστές ως βάρη των συνάψεων, χρησιμοποιούνται για την αποθήκευση της αποκτηθείσας γνώσης.*

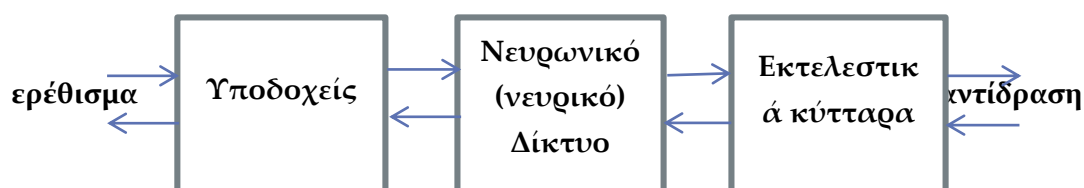
Η θεωρία βασίζεται στη βασική ιδέα της μαθηματικής μοντελοποίησης του ιδιαίτερου τρόπου λειτουργίας του νευρικού συστήματος, στον οποίον οφείλεται η επίτευξη των ανωτέρω επιδόσεων.

Η Έμπνευση από τη Βιολογία

Η αφορμή για την ανάπτυξη της θεωρίας των Νευρωνικών Δικτύων υπήρξε η παρατήρηση πως ο ανθρώπινος εγκέφαλος, ενώ υστερεί έναντι των κλασικών ηλεκτρονικών υπολογιστών σε μεγάλες κατηγορίες προβλημάτων όπως οι αριθμητικοί υπολογισμοί και η διαχείριση συναφών συμβολισμών, εν τούτοις υπερέχει δραματικά στην επίλυση πολύπλοκων προβλημάτων αντίληψης, όπως, ενδεικτικά, η αναγνώριση ανθρώπων από τα χαρακτηριστικά του προσώπου ανάμεσα σε χιλιάδες παρόμοια από μία μόνον θέαση. Ο ανθρώπινος εγκέφαλος είναι οργανωμένος σαν εξαιρετικά πολύπλοκος, μη γραμμικός, παράλληλος υπολογιστής. Έχει τη δυνατότητα να οργανώνει τα στοιχειώδη υπολογιστικά του στοιχεία, του νευρώνες, ώστε αυτοί να επιτελούν υπολογισμούς (π.χ. αναγνώριση μορφών, αντίληψη, κινητικός έλεγχος) κατά πολλές φορές ταχύτερα απ' ότι οι ταχύτεροι τρέχοντες υπολογιστές. Ως παράδειγμα, η λειτουργία της όρασης με την οποία ο εγκέφαλος μας παρέχει αφ' ενός μία αναπαράσταση του περιβάλλοντος γύρω μας, αφ' ετέρου τις πληροφορίες που είναι απαραίτητες για την αλληλεπίδραση με το περιβάλλον, είναι μια διαδικασία επεξεργασίας πληροφοριών. Την εξαιρετικά πολύπλοκη αυτή λειτουργία, ο εγκέφαλος την εκτελεί πολύ ταχύτερα απ' ότι οι υπολογιστές. Ενδεικτικά, η λειτουργία της αναγνώρισης, (πχ της αναγνώρισης ενός γνωστού προσώπου ενσωματωμένου σε ένα άγνωστο υπόβαθρο) γίνεται σε χρόνο 100-200 μs, ενώ λειτουργίες πολύ μικρότερης πολυπλοκότητας είναι πολύ πιο χρονοβόρες κατά την εκτέλεσή τους σε υπολογιστές. Πως επιτυγχάνεται αυτό;

Κατά τη γέννησή του, το ανθρώπινο μυαλό είναι δομημένο και έχει τη δυνατότητα να κατασκευάζει τους δικούς του κανόνες μέσω της διαδικασίας που αποκαλούμε "εμπειρία". Η εμπειρία ενισχύεται με τον χρόνο, με τη σημαντικότερη ανάπτυξη του εγκεφάλου να λαμβάνει χώρα κατά τη διάρκεια των δύο πρώτων ετών από τη γέννηση – παρ' όλα αυτά, η ανάπτυξη συνεχίζεται πολύ πέραν αυτού του ορίου. Η δυνατότητα της "εμπειρίας" συνδέεται άμεσα με την "πλαστικότητα" του εγκεφάλου. Η "πλαστικότητα" σημαίνει την ιδιότητα του νευρικού συστήματος να προσαρμόζεται στο περιβάλλον του. Αυτή την "πλαστικότητα" επιδιώκουν να μιμηθούν, σε μαθηματικό και υπολογιστικό επίπεδο, τα νευρωνικά δίκτυα.

Το ανθρώπινο νευρικό σύστημα μπορεί να θεωρηθεί ως ένα σύστημα τριών σταδίων, όπως απεικονίζεται στο παρακάτω διάγραμμα:



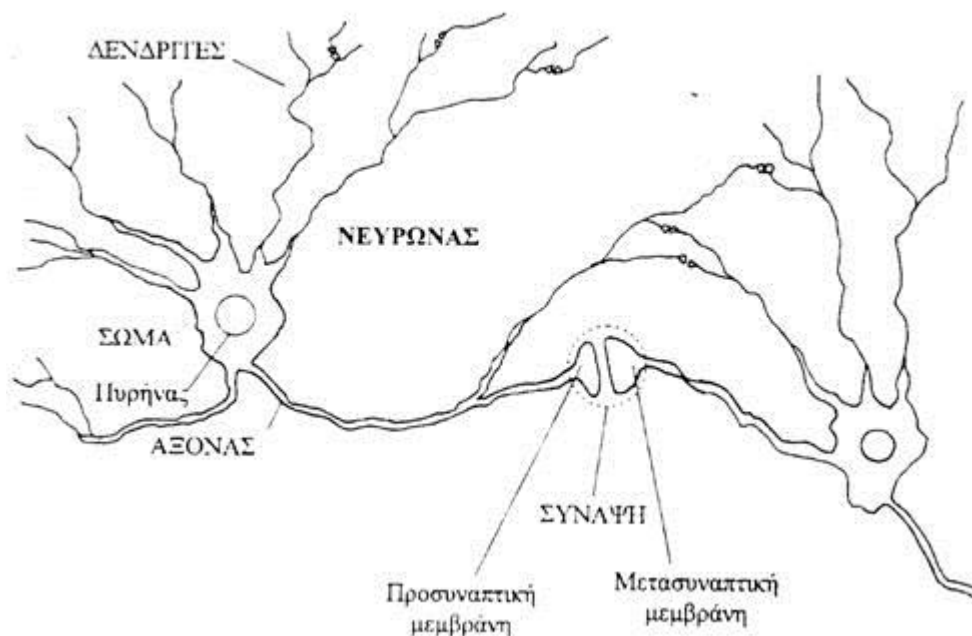
Κεντρικό στοιχείο του συστήματος είναι ο εγκέφαλος, που στο σχήμα απεικονίζεται ως το νευρωνικό (νευρικό) δίκτυο, το οποίο συνεχώς δέχεται πληροφορίες, τις επεξεργάζεται και λαμβάνει τις κατάλληλες αποφάσεις. Στο σχήμα φαίνονται δύο είδη βελών. Αυτά που δείχνουν από αριστερά προς τα δεξιά υποσημαίνουν την προς τα εμπρός μετάδοση σημάτων πληροφοριών μέσα στο σύστημα. Τα βέλη που δείχνουν από δεξιά προς τα αριστερά υποσημαίνουν την ανάδραση εντός του συστήματος. Οι υποδοχείς μετατρέπουν τα ερεθίσματα του ανθρώπινου σώματος και του εξωτερικού περιβάλλοντος σε ηλεκτρικές ωθήσεις που μεταφέρουν πληροφορίες στο νευρωνικό (νευρικό) δίκτυο, δηλαδή στον εγκέφαλο. Τα εκτελεστικά κύτταρα μετατρέπουν τις ηλεκτρικές ωθήσεις που παράγονται από το νευρωνικό δίκτυο σε ευδιάκριτες αντιδράσεις που αποτελούν τα εξαγόμενα του συστήματος.

Η κατανόηση και η μελέτη του συστήματος αυτού έγινε ευχερέστερη όταν το 1911 ο Ramon y Cajal εισήγαγε την έννοια των νευρώνων ως δομικών στοιχείων του εγκεφάλου. Σε γενικές γραμμές, οι νευρώνες είναι πέντε με έξι τάξεις μεγέθους πιο αργοί από τις ηλεκτρονικές ψηφιακές πύλες, όμως ο εγκέφαλος αντισταθμίζει τον σχετικά αργό ρυθμό λειτουργίας με τους τεράστιους αριθμούς νευρώνων (νευρικά κύτταρα) με μαζικές διασυνδέσεις μεταξύ τους. Υπολογίζεται ότι υπάρχουν περίπου 10 δισεκατομμύρια νευρώνες στον φλοιό του ανθρώπινου εγκεφάλου και 60 τρισεκατομμύρια συνάψεις (δηλαδή διασυνδέσεις). Το αποτέλεσμα είναι μια υπολογιστική δομή με εξαιρετικά ισχυρή υπολογιστική ικανότητα.

Οι συνάψεις είναι στοιχειώδεις δομικές και λειτουργικές μονάδες που διαμεσολαβούν μεταξύ των νευρώνων. Το πιο κοινό είδος σύναψης είναι η *χημική*

σύναψη, που λειτουργεί ως εξής: Μία προσυναπτική διαδικασία απελευθερώνει μια "διαβιβαστική" ουσία που διαχέεται κατά μήκος της μεταξύ των νευρώνων και εν συνεργεία επενεργεί σε μία μετασυναπτική διαδικασία. Έτσι, η σύναψη μετατρέπει ένα προσυναπτικό ηλεκτρικό σήμα σε χημικό σήμα και εν συνεχεία σε μετασυναπτικό ηλεκτρικό σήμα.

Η "πλαστικότητα" του εγκεφάλου που αναφέρθηκε προηγουμένως επιτρέπει την ανάπτυξη του νευρικού συστήματος έτσι ώστε να μπορεί να προσαρμοστεί στο περιβάλλον του. Η "πλαστικότητα" αυτή οφείλεται σε δύο μηχανισμούς: στη δημιουργία συναπτικών συνδέσεων μεταξύ των νευρώνων και στην τροποποίηση των υφισταμένων συνάψεων. Οι άξονες, δηλαδή οι γραμμές μεταφοράς, και οι δενδρίτες, οι ζώνες υποδοχής συνιστούν τους δύο τύπους ιών που διακρίνονται μορφολογικά, αφού ένας άξων έχει ομαλότερη επιφάνεια, λιγότερες διακλαδώσεις και μεγαλύτερο μήκος, ενώ ένας δενδρίτης (ονομασία εξ αιτίας της ομοιότητας με δένδρο) έχει ακανόνιστες επιφάνειες και περισσότερες διακλαδώσεις.



Επισημαίνεται ότι στην πραγματικότητα, στον ανθρώπινο εγκέφαλο, οι νευρώνες που αναφέρθηκαν ανωτέρω δεν αποτελούν "ομοιόμορφα" τις βασικές δομές του.

Στην πραγματικότητα, οι νευρώνες οργανώνονται ιεραρχικά σε "δενδριτικές υπομονάδες", και αυτές με τη σειρά τους σε δενδριτικά δένδρα. Τα δενδριτικά δένδρα τα ίδια οργανώνονται ιεραρχικά σε "τοπικά κυκλώματα" και, εν συνεχεία – ιεραρχικά – σε "διαπεριφερειακά κυκλώματα".

Όλη αυτή η πολυπλοκότητα στην οργάνωση των νευρώνων αποτελούν αποκλειστικότητα του εγκεφάλου. Τα νευρωνικά δίκτυα που υλοποιούνται προγραμματιστικά είναι κατά πολύ απλούστερα των φυσικών νευρωνικών δικτύων, αν και βαίνουν συνεχώς πολυπλοκότερα.

Τροπος Λειτουργίας

Η μοντελοποίηση αυτή οφείλεται στις εργασίες των McCulloch και Pitts, ήδη από τη δεκαετία του 1940.

Η μαθηματική προσομοίωση του νευρώνα υπολογίζει το σταθμισμένο άθροισμα n σημάτων εισόδου x_j , $j = 1, 2, \dots, n$ και δημιουργεί μια έξοδο ίση με 1 εάν το άθροισμα αυτό είναι μεγαλύτερο από μία τιμή κατωφλίου u , αλλιώς η έξοδος είναι ίση με 0, δηλαδή:

$$y = \theta \left[\sum_{j=1}^n (w_j x_j - u) \right]$$

όπου $\theta(\cdot)$ η βηματική συνάρτηση και w_j το βάρος της εισόδου j .

Η βασική αυτή προσέγγιση με τον χρόνο γενικεύτηκε, κυρίως με τη χρήση άλλων συναρτήσεων κατωφλίου, πέραν της βηματικής, όπως η σιγμοειδής ή η γκαουσιανή.

Ενώ η παραπάνω μαθηματική προσέγγιση αφορά την προσομοίωση του ενός νευρώνα, η νευρωνική θεωρία επικεντρώνεται στην προσομοίωση του συνολικού μηχανισμού της νευρικής λειτουργίας. Για το σκοπό αυτό, το μαθηματικό μοντέλο του νευρωνικού δικτύου αποτελείται από έναν κατευθυνόμενο γράφο με βάρη, οι

κόμβοι του οποίου αντιστοιχούν στους νευρώνες και οι κατευθυνόμενες ακμές είναι οι συνδέσεις μεταξύ των εισόδων και των εξόδων των νευρώνων.

Ανάλογα με την αρχιτεκτονική του γράφου, τα νευρωνικά δίκτυα διακρίνονται σε δύο βασικές κατηγορίες:

- **Feed-forward δίκτυα**, στα οποία οι γράφοι δεν έχουν βρόχους
- **Δίκτυα με ανατροφοδότηση**, στα οποία εμφανίζονται βρόχοι εξ αιτίας της ανατροφοδότησης

Η βασική διεργασία των νευρωνικών δικτύων – αυτή που ενδιαφέρει στην περίπτωση της Πρόβλεψης – είναι η διαδικασία μάθησης.

Η διαδικασία της μάθησης συνίσταται κατ' ουσίαν στη διαδοχική προσαρμογή της αρχιτεκτονικής του δικτύου – και άρα, των εξισώσεων που το περιγράφουν – κατά τρόπο που να μπορεί να περιγράψει (και να επιτελεί) ικανοποιητικά τη λειτουργία για την οποία προορίζεται. Το δίκτυο "διδάσκεται" τη σωστή αρχιτεκτονική, δηλαδή, στην ουσία, τα βάρη των ακμών του γράφου, με βάση "παραδείγματα", δηλαδή διαθέσιμες πραγματικές περιπτώσεις εφαρμογής της διαδικασίας. Η δυνατότητα να "διδάσκονται" τα νευρωνικά δίκτυα βάσει πραγματικών παραδειγμάτων (που δεν είναι εξ αρχής γνωστά), είναι βασικό τους πλεονέκτημα έναντι κλασικών αναλυτικών μεθόδων.

Η διαδικασία της μάθησης στην πράξη απαιτεί τρία βασικά στοιχεία:

- το μοντέλο του περιβάλλοντος λειτουργίας του νευρωνικού δικτύου, δηλαδή τις διαθέσιμες πληροφορίες – που αποκαλείται "διδασκτικό παράδειγμα"
- τον τρόπο με τον οποίον τα βάρη των ακμών επικαιροποιούνται, δηλαδή τους "διδασκτικούς κανόνες" που διέπουν τη διαδικασία
- τον "διδασκτικό αλγόριθμο", δηλαδή την αλγοριθμική διαδικασία εφαρμογής των διδασκτικών κανόνων

Η θεωρία της μάθησης πρέπει να αντιμετωπίσει τρία βασικά ζητήματα που σχετίζονται με το "διδασκτικό παράδειγμα":

- ✓ Τη δυναμικότητα, δηλαδή το πλήθος των διαθέσιμων παραδειγμάτων που θα αποτελέσουν το εκπαιδευτικό υλικό του δικτύου και που θα μπορούν να αποθηκευτούν και να χρησιμοποιηθούν για τη διαδικασία της μάθησης
- ✓ την πολυπλοκότητα του δείγματος, δηλαδή το πλήθος των παραδειγμάτων που απαιτούνται για την εκπαίδευση του δικτύου, χωρίς αυτό να "υπερπροσδιορίζει" το νευρωνικό δίκτυο στο διδακτικό παράδειγμα και να αποτρέπει την ισχύ του σε άλλα περιπτώσεις
- ✓ την υπολογιστική πολυπλοκότητα, δηλαδή τον απαιτούμενο χρόνο για την εκτέλεση της εκπαίδευσης του νευρωνικού δικτύου με βάση το διδακτικό παράδειγμα.

Τα Βήματα Σχεδίασης Νευρωνικού Δικτύου

Θα παρατεθεί εδώ μια τυπική μέθοδος σχεδίασης 8 βημάτων που προκύπτει από τον συνδυασμό των βημάτων των πλέον διαδεδομένων μεθοδολογιών. Η διαδικασία μπορεί να απαιτήσει την επανάληψη κάποιων βημάτων για τον πλήρη σχεδιασμό ενός δικτύου, κυρίως μεταξύ επιλογής μεταβλητών και διαδικασίας εκπαίδευσης (training).

Τα βήματα της μεθόδου παρατίθενται συνοπτικά σε πίνακα, κι εν συνεχεία παρουσιάζονται αναλυτικά.

1^ο Βήμα	Επιλογή Μεταβλητών
2^ο Βήμα	Συλλογή Δεδομένων
3^ο Βήμα	Αρχική Επεξεργασία Δεδομένων
4^ο Βήμα	Καθορισμός συνόλων Εκπαίδευσης, Δοκιμής και Αξιολόγησης

5^ο Βήμα	Επιλογή βασικών παραμέτρων του δικτύου
6^ο Βήμα	Κριτήρια Αξιολόγησης
7^ο Βήμα	Τελική Προσομοίωση
8^ο Βήμα	Υλοποίηση

Πίνακας 1: Βήματα σχεδίασης νευρωνικού δικτύου

1^ο Βήμα: Επιλογή Μεταβλητών

Ο κρισιμότερος, ίσως, παράγοντας για τον σχεδιασμό του νευρωνικού δικτύου είναι ο καθορισμός των μεταβλητών εισόδου που είναι αναγκαίες για την πρόβλεψη. Η εύστοχη επιλογή των μεταβλητών εισόδου είναι κρίσιμη για την απόδοση του νευρωνικού δικτύου, και αυτή με τη σειρά της εξαρτάται από την κατανόηση του προβλήματος. Καθώς ένα νευρωνικό δίκτυο έχει τη δυνατότητα να ανιχνεύει τις σύνθετες και μη γραμμικές σχέσεις μεταξύ πολλών διαφορετικών μεταβλητών, διαμορφώνονται δείκτες που μπορούν να προκύψουν από τα αρχικά δεδομένα. Από τους δείκτες αυτούς θα διαμορφωθούν οι πραγματικές εισοδοί του νευρωνικού δικτύου.

Η απλούστερη μέθοδος χρησιμοποιεί τις αρχικές τιμές της εξαρτώμενης μεταβλητής ή της πρώτης διαφοράς της ως εισόδους. Τέτοια μοντέλα έχουν καταλήξει σε ακριβέστερα συμπεράσματα από τις παραδοσιακές μεθόδους Box-Jenkins. Ακόμη συνηθέστερη προσέγγιση είναι να υπολογίζονται οι τεχνικοί δείκτες που είναι βασισμένοι μόνο στις προηγούμενες παρατηρήσεις.

2^ο Βήμα: Συλλογή Δεδομένων

Στη φάση αυτή εξετάζεται το κόστος και η διαθεσιμότητα των επιλεγμένων μεταβλητών του προηγούμενου βήματος. Τα δεδομένα είναι εύκολα διαθέσιμα από πολλούς προμηθευτές - οι βασικές πληροφορίες είναι συχνά δυσκολότερο να

αποκτηθούν. Ο χρόνος που ξοδεύεται κατά τη συλλογή των στοιχείων δεν μπορεί να χρησιμοποιηθεί για την προεπεξεργασία, την εκπαίδευση και την αξιολόγηση της επίδοσης των νευρωνικών δικτύων. Τα στοιχεία θα πρέπει να ελεγχθούν για λάθη, εξετάζοντας αλλαγές, αποκλίσεις ή ακόμα απώλειες παρατηρήσεων.

Οι ελλείπουσες παρατηρήσεις αντιμετωπίζονται με διάφορους τρόπους. Ένας τρόπος είναι να απορριφθούν όλες ή να υποτεθεί ότι παραμένουν το ίδιο με την παρεμβολή ή τον υπολογισμό του μέσου όρου από τις κοντινές τιμές. Ένας άλλος δυνατός τρόπος είναι η κωδικοποίηση ενός νευρώνα εισόδου που λαμβάνει την τιμή 1 όταν ελλείπει τιμή και την τιμή 0 όταν υφίσταται.

3^ο Βήμα: Αρχική Επεξεργασία Δεδομένων

Στη φάση αυτή γίνεται ανάλυση και μετασχηματισμός των μεταβλητών εισόδου και εξόδου. Ο στόχος είναι η ελαχιστοποίηση του θορύβου, η αναγνώριση και ανάδειξη σημαντικών σχέσεων, η ανίχνευση των τάσεων και η εξομάλυνση της κατανομής της μεταβλητής ώστε να βοηθήσει το νευρωνικό δίκτυο στην εκμάθηση αυτών. Καθώς τα νευρωνικά δίκτυα εκπαιδεύονται μιμούμενα πρότυπα, η παρουσίαση των δεδομένων αποτελεί κρίσιμο παράγοντα στην σχεδίαση ενός επιτυχημένου νευρωνικού δικτύου. Κατ' ελάχιστον θα πρέπει τα ακατέργαστα δεδομένα να κλιμακωθούν μεταξύ των ανώτερων και κατώτερων ορίων των συναρτήσεων μεταφοράς .

Δύο κοινοί μετασχηματισμοί δεδομένων τόσο στις παραδοσιακές μεθόδους πρόβλεψης όσο και στην πρόβλεψη με νευρωνικά δίκτυα, είναι αρχικά η χρησιμοποίηση των διαφορών μεταξύ των διαφορετικών τιμών που παίρνουν μια μεταβλητή και η επιλογή του φυσικού λογαρίθμου αυτής. Το αποτέλεσμα είναι η αφαίρεση της γραμμικής τάσης από τα δεδομένα. Ο λογαριθμικός μετασχηματισμός είναι χρήσιμος για δεδομένα που μπορούν να πάρουν και μικρές και μεγάλες τιμές και χαρακτηρίζεται από μια εκτεταμένη δεξιά διανομή ουρών.

Ένας άλλος δημοφιλής μετασχηματισμός δεδομένων είναι να χρησιμοποιηθούν οι αναλογίες των μεταβλητών εισόδου. Οι αναλογίες δίνουν έμφαση σε σημαντικές σχέσεις ενώ συγχρόνως συντηρούν τους βαθμούς ελευθερίας επειδή λιγότεροι

νευρώνες εισόδου χρειάζονται για να κωδικοποιήσουν τις ανεξάρτητες μεταβλητές.

Η εξομάλυνση των δεδομένων εισόδου και εξόδου, με τη χρησιμοποίηση είτε της απλής είτε της εκθετικής εξομάλυνσης, υιοθετείται συχνά. Μερικές φορές βέβαια είναι προτιμότερο να χρησιμοποιούνται οι κινητοί μέσοι όροι για την εξομάλυνση των ανεξάρτητων μεταβλητών και την πρόβλεψη τάσεων.

Η δειγματοληψία ή το φιλτράρισμα των δεδομένων αναφέρονται στην αφαίρεση κάποιων παρατηρήσεων από τα training και test sets για να δημιουργήσουν μια πιο ομοιόμορφη διανομή. Ο τύπος φιλτραρίσματος που χρησιμοποιείται πρέπει να είναι σύμφωνος με τους αντικειμενικούς στόχους του ερευνητή. Το δίκτυο ελαχιστοποιεί το ποσό των τετραγωνικών σφαλμάτων (ή άλλων εξισώσεων σφαλμάτων) σε όλες τις μεθόδους εκπαίδευσης. Ο ερευνητής πρέπει να είναι ξεκάθαρος, στο τι ακριβώς πρέπει να μάθει το νευρωνικό δίκτυο. Ένα άλλο πλεονέκτημα του φιλτραρίσματος είναι η μείωση των γεγονότων εκπαίδευσης, κάτι που επιτρέπει την εξέταση περισσότερων μεταβλητών εισόδου, την τυχαία επιλογή των βαρών ή/και των κρυμμένων επιπέδων από την περίπτωση εκπαίδευσης μεγάλων συνόλων δεδομένων.

Στην πράξη γίνονται πολλές δοκιμές. Η πιο κλασική μέθοδος επιλογής των κατάλληλων μεταβλητών εισόδου είναι η δοκιμή διαφόρων συνδυασμών. Αν και μια τέτοια διαδικασία είναι αρκετά χρονοβόρα και απαιτεί μεγάλη υπολογιστική πολυπλοκότητα, αναγνωρίζει τις περιπτώσεις ότι κάποιες μεταβλητές μπορεί να είναι ιδανικές για την πρόβλεψη μόνο όταν συνδυάζονται με άλλες μεταβλητές.

4^ο Βήμα: Καθορισμός συνόλων Εκπαίδευσης, Δοκιμής και Αξιολόγησης

Η κοινή πρακτική είναι να διαιρεθούν τα δεδομένα της κάθε χρονοσειράς στα τρία παραπάνω σύνολα. Το **σύνολο εκπαίδευσης** είναι το μεγαλύτερο σύνολο και

χρησιμοποιείται από το νευρωνικό δίκτυο για να μάθει τη μορφή που έχουν τα δεδομένα. Συνήθως αποτελεί το 60-80% του συνόλου των δεδομένων. Το **σύνολο δοκιμής**, που κυμαίνεται μεταξύ 10-30% του συνόλου των δεδομένων, χρησιμοποιείται για να αξιολογήσει τη δυνατότητα γενίκευσης ενός υποθετικά εκπαιδευμένου δικτύου. Ο ερευνητής θα επιλέξει το δίκτυο που αποδίδει καλύτερα στο **σύνολο δοκιμής**. Ένας τελικός έλεγχος στην απόδοση του εκπαιδευμένου δικτύου γίνεται χρησιμοποιώντας το **σύνολο αξιολόγησης**, που είναι συνήθως 10% του συνόλου. Το **σύνολο αξιολόγησης** πρέπει να αποτελείται από τις πιο πρόσφατες παρακείμενες παρατηρήσεις.

Το σύνολο δοκιμής μπορεί είτε να επιλεγεί τυχαία από το σύνολο εκπαίδευσης είτε να αποτελείται από ένα σύνολο παρατηρήσεων αμέσως μετά από το σύνολο δοκιμής. Το πλεονέκτημα της τυχαίας επιλογής είναι ότι αποφεύγεται ο κίνδυνος του να χαρακτηριστεί το σύνολο δοκιμής από ένα μόνο είδος των δεδομένων. Αντιθέτως το πλεονέκτημα της χρησιμοποίησης των παρατηρήσεων που ακολουθούν το σύνολο εκπαίδευσης ως σύνολο δοκιμής είναι ότι αυτές είναι οι πιο πρόσφατες παρατηρήσεις που μπορεί αν είναι σημαντικότερες από τα παλαιότερα δεδομένα.

Τα τυχαία επιλεγμένα δεδομένα του συνόλου δοκιμής δεν θα πρέπει να αντικαθιστώνται στο σύνολο εκπαίδευσης γιατί αυτό επηρεάζει την ικανότητα γενίκευσης. Επιπλέον μια ντετερμινιστική μέθοδος, όπως το να επιλέγεται κάθε ν-οστή παρατήρηση για στοιχείο του συνόλου δοκιμής, δεν προτείνεται για τον απλό λόγο ότι μπορεί να οδηγήσει σε κύκλους των δεδομένων επί των οποίων έγινε δειγματοληψία.

Προτείνεται να κλιμακοποιούνται μαζί τα σύνολα εκπαίδευσης και δοκιμών, εφόσον και ο σκοπός του συνόλου δοκιμών είναι ο καθορισμός της ικανότητας γενίκευσης του δικτύου. Εντούτοις, με κανένα τρόπο, δεν πρέπει το σύνολο αξιολόγησης να κλιμακοποιείται μαζί είτε με το σύνολο εκπαίδευσης είτε με το

σύνολο δοκιμής γιατί κάτι τέτοιο θα οδηγήσει σε με ασφαλή ακεραιότητα του συνόλου αξιολόγησης το οποίο αποτελεί και τον τελικό και ανεξάρτητο έλεγχο στο νευρωνικό δίκτυο. Στην πραγματικότητα, ο ερευνητής δεν μπορεί με κανέναν τρόπο να γνωρίζει την ακριβή σειρά των μελλοντικών τιμών, αλλά να κάνει μόνο μια λογική εκτίμηση στη σειρά του συνόλου εκπαίδευσης και συνόλου δοκιμής.

Βήμα 5ο : Επιλογή βασικών παραμέτρων του δικτύου

Υπάρχει ένας άπειρος αριθμός τρόπων να κατασκευαστεί ένα νευρωνικό δίκτυο. Η αρχιτεκτονική του δικτύου, οι ιδιότητες κάθε μεμονωμένου νευρώνα, η λειτουργία μεταφοράς του και το πώς οι είσοδοι συνδυάζονται καθορίζουν το είδος του νευρωνικού δικτύου. Η αρχιτεκτονική ενός νευρωνικού δικτύου καθορίζει τη δομή του, συμπεριλαμβανομένου του αριθμού των νευρώνων σε κάθε επίπεδο και του αριθμού και του τύπου των διασυνδέσεων.

Ο αριθμός νευρώνων εισόδου είναι μια από τις ευκολότερες παραμέτρους για να επιλέγουν, αφού πρώτα υποστούν επεξεργασία οι ανεξάρτητες μεταβλητές, επειδή κάθε ανεξάρτητη μεταβλητή αντιπροσωπεύεται και τον δικό της νευρώνα εισόδου.

Παρακάτω θα εξετάσουμε την επιλογή του αριθμού των κρυφών στρώματων, των κρυφών νευρώνων κάθε στρώματος, των νευρώνων εξόδου και των συναρτήσεων μεταφοράς.

5.1 Αριθμός κρυμμένων επιπέδων

Τα κρυμμένα επίπεδα (στρώματα) παρέχουν στο δίκτυο, τη δυνατότητα να γενικευτεί. Ένα νευρωνικό δίκτυο με ένα κρυμμένο επίπεδο και με ικανοποιητικό αριθμό κρυφών νευρώνων είναι σε θέση να προσεγγίζει οποιαδήποτε λειτουργία. Στην πράξη, τα νευρωνικά δίκτυα με ένα και περιστασιακά δύο κρυμμένα επίπεδα χρησιμοποιούνται ευρέως και έχουν αποδώσει πολύ καλά. Η αύξηση του αριθμού

των κρυφών στρωμάτων αυξάνει επίσης το χρόνο υπολογισμού και τον κίνδυνο "υπερπροσαρμογής" που οδηγεί σε κακές προβλέψεις. Η "υπερπροσαρμογή" εμφανίζεται όταν ένα μοντέλο πρόβλεψης έχει πολύ λίγους βαθμούς ελευθερίας. Με άλλα λόγια, όταν έχει σχετικά λίγες παρατηρήσεις σε σχέση με τις παραμέτρους του και επομένως είναι σε θέση να απομνημονεύσει τα μεμονωμένα σημεία παρά να μάθει τη γενική μορφή.

Στην περίπτωση των νευρωνικών δικτύων, ο αριθμός των βαρών, που συνδέεται άμεσα με τον αριθμό των κρυμμένων στρωμάτων και νευρώνων καθώς και το μέγεθος του συνόλου εκπαίδευσης, καθορίζουν την πιθανότητα υπερπροσαρμογής. Όσο μεγαλύτερος ο αριθμός των βαρών σε σχέση με το μέγεθος του συνόλου εκπαίδευσης, τόσο μεγαλύτερη είναι η δυνατότητα του δικτύου να συγκρατήσει την ιδιοσυγκρασία των μεμονωμένων παρατηρήσεων. Αυτό έχει ως συνέπεια, η γενίκευση για το σύνολο αξιολόγησης να χάνεται και το μοντέλο να μη χρησιμοποιείται σε μια πραγματική πρόβλεψη. Επομένως, συνιστάται όλα τα νευρωνικά δίκτυα να αρχίζουν με ένα ή το πολύ δύο κρυμμένα επίπεδα. Εάν ένα νευρωνικό δίκτυο τεσσάρων στρωμάτων (με δύο κρυφά στρώματα) αποδειχτεί ανεπαρκές αφού έχει εξετάσει πολλαπλούς κρυμμένους νευρώνες χρησιμοποιώντας έναν λογικό αριθμό τυχαία επιλεγμένων αρχικών βαρών, στη συνέχεια ο ερευνητής πρέπει να τροποποιήσει μερικές φορές τις μεταβλητές εισόδου πριν προσθέσει ένα τρίτο κρυφό επίπεδο. Σύμφωνα με τη θεωρία και όλη την εμπειρική εργασία έως σήμερα, τα δίκτυα με περισσότερα από τέσσερα στρώματα δε βελτιώνουν τα αποτελέσματα.

5.2 Αριθμός κρυμμένων νευρώνων

Ο βέλτιστος αριθμός κρυφών νευρώνων δε μπορεί να υπολογιστεί αναλυτικά, παρά την κεντρική σημασία του για την κατασκευή ενός νευρωνικού δικτύου. Ο προσδιορισμός τους στην πράξη γίνεται πειραματικά. Ερευνητικά, έχουν προταθεί αρκετές εμπειρικές μέθοδοι για τον καθορισμό του βέλτιστου αριθμού, με απλούστερη αλλά ικανοποιητική την προσέγγιση του Masters, κατά την οποία,

για ένα δίκτυο τριών στρωμάτων με n νευρώνες εισόδου και m νευρώνες εξόδου, το κρυμμένο επίπεδο θα έχει \sqrt{nm} νευρώνες. Ο πραγματικός αριθμός κρυφών νευρώνων μπορεί ακόμα να κυμανθεί από το $1/2$ ως 2 φορές την τιμή που προκύπτει από τον κανόνα αυτό ανάλογα με τη πολυπλοκότητα του προβλήματος. Σε κάθε περίπτωση, είναι σημαντικό να σημειωθεί ότι οι κανόνες που υπολογίζουν τον αριθμό κρυφών νευρώνων ως πολλαπλάσιο του αριθμού νευρώνων εισόδου, υποθέτουν ότι το σύνολο εκπαίδευσης είναι τουλάχιστον δύο φορές μεγαλύτερο από τον αριθμό των βαρών και κατά προτίμηση τέσσερις ή περισσότερες καθώς αυξάνεται. Εάν αυτό δεν ισχύει, τότε οι εμπειρικές αυτές μέθοδοι οδηγούν σε υπερπροσαρμοσμένα μοντέλα, δεδομένου ότι ο αριθμός κρυφών νευρώνων εξαρτάται άμεσα από τον αριθμό νευρώνων εισόδου (που καθορίζουν στη συνέχεια τον αριθμό των βαρών). Η λύση είναι είτε να αυξηθεί το μέγεθος του συνόλου εκπαίδευσης είτε, αν αυτό δεν είναι εφικτό, να τεθεί ένα ανώτερο όριο στον αριθμό των νευρώνων εισόδου έτσι ώστε ο αριθμός βαρών είναι τουλάχιστον ο μισός από τον αριθμό των γεγονότων του συνόλου εκπαίδευσης. Η επιλογή των μεταβλητών εισόδου γίνεται περισσότερο κρίσιμη σε μικρά δίκτυα από τη στιγμή που δεν υπάρχει η πολυτέλεια της παρουσίας στο δίκτυο ενός μεγάλου αριθμού εισόδων έτσι ώστε να μπορεί τότε αυτό με τη σειρά του να αγνοεί τις μη χρήσιμες εισόδους.

Η επιλογή του σωστού αριθμού κρυφών νευρώνων περιλαμβάνει αρκετό πειραματισμό. Οι τρεις επικρατέστερες μέθοδοι είναι οι :

- α) σταθερή,
- β) κατασκευαστική
- γ) καταστροφική.

Στη πρώτη προσέγγιση, μια ομάδα νευρωνικών δικτύων με διαφορετικούς αριθμούς κρυφών νευρώνων εκπαιδεύεται και καθεμία αξιολογείται στο σύνολο

εκπαίδευσης, χρησιμοποιώντας έναν λογικό αριθμό τυχαία επιλεγμένων αρχικών βαρών. Η αύξηση στον αριθμό κρυφών νευρώνων μπορεί να είναι ένας, δύο ή περισσότεροι, ανάλογα με τους διαθέσιμους υπολογιστικούς πόρους. Το δίκτυο με το μικρότερο σφάλμα επιλέγεται επειδή είναι σε θέση να γενικεύσει καλύτερα. Αυτή η προσέγγιση είναι χρονοβόρα, αλλά γενικά λειτουργεί πολύ καλά.

Η κατασκευαστική και η καταστροφική προσέγγιση περιλαμβάνουν την αλλαγή του αριθμού των κρυφών νευρώνων κατά τη διάρκεια της εκπαίδευσης αντί της δημιουργίας χωριστών δικτύων, καθένα με διαφορετικό αριθμό κρυφών νευρώνων, όπως στην πρώτη προσέγγιση. Πολλά εμπορικά νευρωνικά πακέτα λογισμικού δικτύων δεν υποστηρίζουν αυτές τις μεθόδους. Η κατασκευαστική προσέγγιση περιλαμβάνει την προσθήκη κρυφών νευρώνων έως ότου αρχίσει η χειροτέρευση του δικτύου. Η καταστροφική προσέγγιση είναι παρόμοια, εκτός από το ότι οι κρυφοί νευρώνες αφαιρούνται κατά τη διάρκεια της εκπαίδευσης.

Γενικά ο κανόνας είναι να επιλέγεται πάντα το δίκτυο που αποδίδει καλύτερα στο σύνολο ελέγχου με το λιγότερο αριθμό κρυμμένων νευρώνων. Κατά τις δοκιμές του αριθμού των νευρώνων είναι σημαντικό να κρατηθούν όλες οι άλλες παράμετροι σταθερές. Η αλλαγή οποιασδήποτε παραμέτρου δημιουργεί ουσιαστικά ένα νέο νευρωνικό δίκτυο με διαφορετικό εύρος σφάλματος το οποίο με τη σειρά του επηρεάζει την επιλογή του βέλτιστου αριθμού των κρυφών νευρώνων.

5.3 Αριθμός νευρώνων εξόδου

Ο καθορισμός του αριθμού νευρώνων εξόδου είναι απλός, καθώς υπάρχουν συγκεκριμένοι λόγοι να χρησιμοποιείται πάντα ένας μόνο νευρώνας εξόδου. Τα νευρωνικά δίκτυα με πολλές εξόδους, ειδικά εάν αυτές οι εξοδοί χωρίζονται ευρέως, θα παράγουν χειρότερα αποτελέσματα σε σύγκριση με ένα δίκτυο με μια μοναδική έξοδο. Ένα νευρωνικό δίκτυο εκπαιδεύεται με την επιλογή των βαρών έτσι ώστε ο μέσο όρος σφαλμάτων σε όλους τους νευρώνες εξόδου να ελαχιστοποιείται.

5.4 Συναρτήσεις μεταφοράς

Οι συναρτήσεις μεταφοράς είναι μαθηματικοί τύποι που καθορίζουν την έξοδο ενός νευρώνα επεξεργασίας και είναι γνωστές επίσης ως εξισώσεις μετασχηματισμού, συμπίεσης και ενεργοποίησης. Η πλειοψηφία των περισσότερων μοντέλων νευρωνικών δικτύων χρησιμοποιούν τη σιγμοειδή συνάρτηση, αλλά και άλλες όπως την υπερβολική εφαπτομένη, τη βηματική και τη γραμμική. Ο σκοπός της

συνάρτησης μεταφοράς είναι να μη φτάσουν οι έξοδοι σε πολύ μεγάλες τιμές, κάτι που μπορεί να παραλύσει τα νευρωνικά δίκτυα και με αυτόν τον τρόπο να εμποδίσει τη διαδικασία εκπαίδευσης.

Οι γραμμικές συναρτήσεις μεταφοράς δεν είναι χρήσιμες στην περίπτωση της μη γραμμικότητας. Οι συναρτήσεις μεταφοράς όπως οι σιγμοειδείς χρησιμοποιούνται συνήθως για χρονοσειρές επειδή είναι μη γραμμικές και συνεχώς διαφοροποιήσιμες, ιδιότητες που είναι επιθυμητές για την εκμάθηση των δικτύων.

Ο Klimasauskas σε σχετική μελέτη τόνισε ότι σε περιπτώσεις όπου ένα δίκτυο πρόκειται να μάθει από την μέση συμπεριφορά, τότε θα πρέπει να χρησιμοποιηθεί μια Sigmoid συνάρτηση μεταφοράς. Σε περιπτώσεις όπου η εκμάθηση του δικτύου περιλαμβάνει αποκλίσεις από τη μέση τιμή, τότε θα πρέπει να χρησιμοποιηθεί η συνάρτηση υπερβολικής εφαπτομένης. Η κλιμακωτή και η βηματική συνάρτηση προτείνονται για δυαδικές μεταβλητές εφόσον η σιγμοειδής συνάρτηση μεταφοράς πλησιάζει ασυμπτωτικά τα μηδέν και ένα. Σε ένα κλασσικό νευρωνικό δίκτυο που χρησιμοποιεί τον αλγόριθμο back-propagation (BP network), οι νευρώνες εισόδου χρησιμοποιούν συνήθως γραμμικές συναρτήσεις μεταφοράς ενώ όλοι οι άλλοι νευρώνες του δικτύου χρησιμοποιούν σιγμοειδή συνάρτηση.

Η απλή γραμμική κλιμάκωση είναι δεκτική στις απότομες αλλαγές με παροδικό χαρακτήρα διότι δεν αλλάζει την ομοιομορφία της διανομής παρά μόνο κλιμακώνει μερικά τις αλλαγές αυτές μέσα στα κατάλληλα όρια της συνάρτησης μεταφοράς. Επισημαίνεται ότι τα περισσότερα πακέτα λογισμικού νευρωνικών δικτύων κλιμακώνουν αυτόματα όλες τις μεταβλητές μέσα στο κατάλληλο διάστημα.

6ο Βήμα: Κριτήρια αξιολόγησης

Η πιο συνήθης συνάρτηση σφάλματος που ελαχιστοποιείται στα νευρωνικά δίκτυα είναι το άθροισμα των τετραγωνικών σφαλμάτων. Άλλες λειτουργίες σφάλματος που προσφέρονται από τους προμηθευτές λογισμικού περιλαμβάνουν τις ελάχιστες απόλυτες αποκλίσεις, τις ασυμμετρικές ελαχίστων τετραγώνων και τις διαφορές ποσοστών. Αυτές οι λειτουργίες σφάλματος μπορούν να μην είναι τα τελικά κριτήρια αξιολόγησης, δεδομένου ότι άλλες μέθοδοι αξιολόγησης της πρόβλεψης, όπως το MAPE, δεν ελαχιστοποιείται στα νευρωνικά δίκτυα.

7ο Βήμα: Τελική Προσομοίωση

Η εκπαίδευση ενός νευρωνικού δικτύου για την εκμάθηση του προτύπου που πρέπει να ακολουθηθεί περιλαμβάνει την επαναληπτική παρουσίασή του με παραδείγματα ήδη γνωστών απαντήσεων. Ο στόχος της εκπαίδευσης είναι να βρεθεί το σύνολο των βαρών μεταξύ των νευρώνων που καθορίζουν το συνολικό ελάχιστο της συνάρτησης σφάλματος. Από τη στιγμή που το μοντέλο δεν είναι υπερπροσαρμοσμένο, αυτό το σύνολο των βαρών πρέπει να παρέχει καλή γενίκευση. Το BP νευρωνικό δίκτυο χρησιμοποιεί έναν αλγόριθμο κλιμακωτής καθόδου ο οποίος ρυθμίζει τα βάρη έτσι ώστε να ακολουθούν την πιο απότομη κλίση της επιφάνειας της συνάρτησης σφάλματος. Βέβαια η εύρεση του ολικού ελαχίστου δεν είναι πάντα σίγουρη μιας και η επιφάνεια της συνάρτησης

σφάλματος περιέχει πολλά τοπικά ελάχιστα στα οποία μπορεί ο αλγόριθμος να 'κολλήσει'. Σε αυτή την ενότητα θα συζητήσουμε το πότε πρέπει να σταματάει η εκπαίδευση ενός νευρωνικού δικτύου καθώς και την επιλογή του ρυθμού εκμάθησης και τιμών της ταχύτητας.

7.1 Αριθμός Επαναλήψεων

Υπάρχουν 2 απόψεις σχετικά με το σημείο στο οποίο πρέπει η εκπαίδευση να σταματήσει. Η πρώτη τονίζει τον κίνδυνο παγίδευσης σε ένα τοπικό ελάχιστο και τη δυσκολία του να φτάσουμε σε ένα ολικό ελάχιστο. Ο ερευνητής πρέπει μόνο να σταματήσει όταν δεν υπάρχει πια καμία βελτίωση στη συνάρτηση σφάλματος, βασισμένη σε ένα λογικό αριθμό τυχαία επιλεγμένων αρχικών βαρών. Το σημείο στο οποίο το δίκτυο δεν βελτιώνεται καλείται σύγκλιση. Η δεύτερη άποψη υποστηρίζει μια σειρά από διακοπές εκπαίδευσης-ελέγχου. Η εκπαίδευση του δικτύου σταματά μετά από έναν προκαθορισμένο αριθμό επαναλήψεων, εκτιμάται η δυνατότητα του δικτύου να γενικεύσει στο σύνολο δοκιμών και η εκπαίδευση συνεχίζεται. Η γενίκευση είναι η ιδέα ότι ένα μοντέλο βασισμένο σε ένα δείγμα δεδομένων είναι κατάλληλο για την πρόβλεψη του συνόλου. Το δίκτυο για το οποίο το σφάλμα του συνόλου δοκιμής είναι πολύ χαμηλό, επιλέγεται δεδομένου ότι γενικεύει καλύτερα. Η επίκριση της διαδικασίας εκπαίδευσης-ελέγχου είναι ότι επιπρόσθετες διακοπές θα μπορούσαν να προκαλέσουν μια περαιτέρω πτώση του σφάλματος προτού αυξηθεί ξανά ή ακόμα και να μειωθεί το σφάλμα ασυμπτωτικά.

Και οι δυο απόψεις συμφωνούν ότι η γενίκευση είναι ο απώτερος στόχος και χρησιμοποιούν σύνολα ελέγχου για να αξιολογήσουν έναν μεγάλο αριθμό δικτύων. Το σημείο στο οποίο αυτές οι δύο προσεγγίσεις αρχίζουν να διαχωρίζονται εστιάζεται στην ιδέα του υπερ-εκπαίδευσης σε σύγκριση με την υπερπροσαρμογή. Η προσέγγιση σύγκλισης δηλώνει ότι δεν υπάρχει κανένα σημείο υπερ-εκπαίδευσης αλλά μόνον υπερπροσαρμογής, το οποίο είναι απλά ένα χαρακτηριστικό ενός δικτύου που έχει πάρα πολλά βάρη. Η λύση είναι να μειωθεί

ο αριθμός κρυμμένων νευρώνων (ή των κρυμμένων επιπέδων αν υπάρχουν περισσότερα του ενός) ή/και να αυξηθεί το μέγεθος του συνόλου εκπαίδευσης. Η προσέγγιση εκπαίδευσης-δοκιμής επιχειρεί να διαφυλάξει το δίκτυο από την υπερπροσαρμογή σταματώντας την ; διαδικασία της εκπαίδευσης (training) βασιζόμενη στην ικανότητα του δικτύου να γενικεύεται.

Το πλεονέκτημα της προσέγγισης σύγκλισης είναι ότι κάποιος μπορεί αν είναι αισιόδοξος ότι το ολικό ελάχιστο έχει επιτευχθεί. Το ίδιο πράγμα για την προσέγγιση train-test είναι δυσκολότερο να υποτεθεί γιατί τα αρχικά βάρη επιλέγονται τυχαία και η μέση συσχέτιση μπορεί να έχει μεγάλη διακύμανση κατά τη διάρκεια της εκπαίδευσης. Ένα άλλο πλεονέκτημα είναι ότι ο ερευνητής έχει δύο λιγότερες παραμέτρους για να ασχολείται, συγκεκριμένα με το σημείο στο οποίο θα σταματήσει η εκπαίδευση και με τη μέθοδο εκτίμησης για το ποιο από τα δίκτυα εκπαίδευσης-ελέγχου είναι το ιδανικό. Ένα πλεονέκτημα της προσέγγισης εκπαίδευσης-ελέγχου είναι ότι δίκτυα με λιγότερους βαθμούς ελευθερίας μπορούν να εφαρμοστούν με καλύτερη γενίκευση από την προσέγγιση σύγκλισης η οποία θα κατέληγε σε υπερπροσαρμογή. Επιπλέον, η προσέγγιση εκπαίδευσης-ελέγχου απαιτεί μικρότερο χρόνο εκπαίδευσης.

Για να καθοριστεί ο μέγιστος αριθμός επαναλήψεων του δικτύου πρέπει είτε να προσδιοριστεί κάποιο μέτρο σφάλματος για κάθε επανάληψη (ενδεικτικά: η μέση συσχέτιση, το άθροισμα των τετραγωνικών σφαλμάτων, ή κάποιο άλλο κατάλληλο μέτρο σφάλματος για κάθε επανάληψη) ή να προκαθοριστούν διαστήματα μέχρι το σημείο όπου η βελτίωση είναι αμελητέα (συνήθως μέχρι ένα μέγιστο 10.000 επαναλήψεων). Κάθε επανάληψη μπορεί να σχεδιαστεί εύκολα εάν το λογισμικό νευρωνικού δικτύου δημιουργεί ένα αρχείο στατιστικών ή εάν αυτό δεν ισχύει, η μέση συσχέτιση μπορεί να καταγραφεί σε διαστήματα των 100 ή 200 από την οθόνη του υπολογιστή. Αφού σχεδιάσει τη μέση συσχέτιση για διάφορα τυχαία επιλεγμένα αρχικά βάρη, ο ερευνητής μπορεί να επιλέξει το μέγιστο αριθμό επαναλήψεων, βασισμένων στο σημείο όπου η μέση συσχέτιση σταματά να αυξάνεται γρήγορα και ισιώνει.

5. Η νευρωνική εκδοχή της Μεθόδου Θήτα

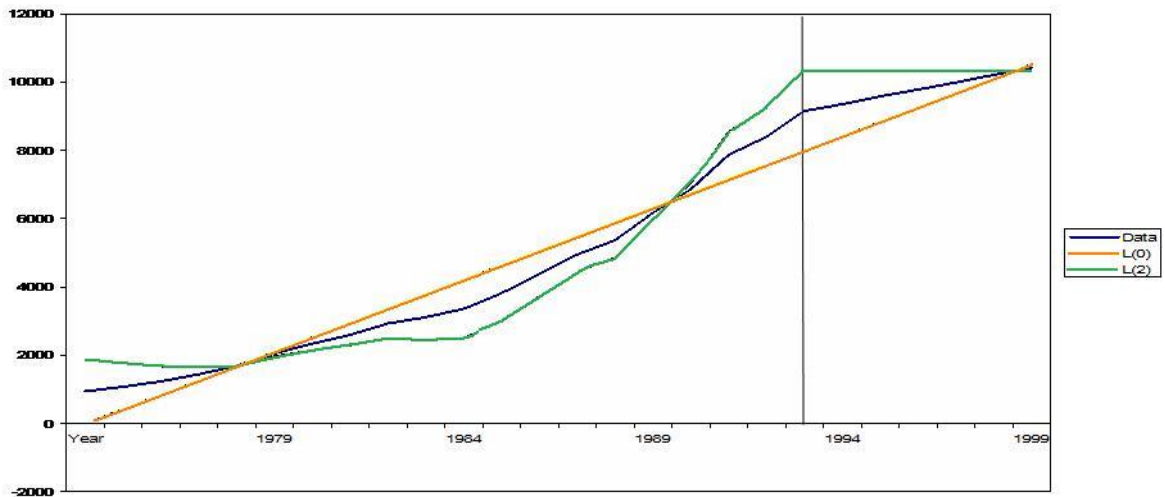
Νευρωνικά Δίκτυα και Πρόβλεψη

Μία από τις εφαρμογές των τεχνητών νευρωνικών δικτύων είναι και η χρήση τους στην υλοποίηση μεθόδων προβλέψεως. Στη βιβλιογραφία υπάρχουν πολλές αναφορές στην επίδοση νευρωνικών δικτύων στην πρόγνωση, τα οποία είναι συχνά αντιφατικά οδηγούν σε άτοπα συμπεράσματα. Το βασικός λόγος γι' αυτό είναι ότι υπάρχει ένας μεγάλος αριθμός παραγόντων που μπορεί να επηρεάσουν την ικανότητα ενός μοντέλου να παραγάγει εύστοχες και ακριβείς προγνώσεις, όπως ενδεικτικά η αρχιτεκτονική του δικτύου, η μέθοδος εκπαίδευσης, τα αρχικά σύνολα δεδομένων. Σε μερικές περιπτώσεις, τα νευρωνικά δίκτυα αποδίδουν χειρότερα αποτελέσματα έναντι των γραμμικών στατιστικών μεθόδων. Ο λόγος για αυτό μπορεί να είναι απλώς ότι τα δεδομένα είναι γραμμικά, χωρίς ιδιαίτερα μεγάλες διακυμάνσεις. Στην περίπτωση αυτή, αναμένεται τα γραμμικά μοντέλα να αποδίδουν καλύτερα αποτελέσματα έναντι των νευρωνικών δικτύων

Η Μεθοδος Theta AI.

Η βασική ιδέα της προσέγγισης Theta AI είναι η ενίσχυση της απλής Μεθόδου Θ με έναν μηχανισμό καταμερισμού των βαρών των βοηθητικών γραμμών θ που χρησιμοποιούνται στην μέθοδο. Η βασική μορφή της Μεθόδου θ χρησιμοποιεί τις γραμμές LRL (γραμμική παλινδρόμηση) ή αλλιώς $L(\theta=0)$ και τη γραμμή $L(\theta=2)$. Για την παραγωγή της τελικής πρόβλεψης, η LRL προεκτείνεται περαιτέρω στα σημεία του οριζοντα πρόβλεψης και στην $L(\theta=2)$ εφαρμόζεται μέθοδος SES (Simple Exponential Smoothing – Εκθετική Εξομάλυνση Σταθερού Επιπέδου) ή Naive μέθοδος. Ο συνδυασμός των προβλέψεων των γραμμών $L(\theta=0)$ και $L(\theta=2)$, με βάρος 50% η κάθε μια, δίνει το τελικό αποτέλεσμα.

Figure 1. M3 Competition - N0001YB001-YEARLY/MCRO



Η μέθοδος Theta AI διαφοροποιείται στην απόδοση των τελικών βαρών της μεθόδου. Συγκεκριμένα, τα βάρη καθορίζονται από ένα νευρωνικό δίκτυο που έχει εκπαιδευτεί με τα δεδομένα των γραμμών $L(\theta=0)$ και $L(\theta=2)$.

Ακολουθεί μια συνοπτική παρουσίαση των βημάτων της Theta AI:

1. Αποεποχικοποίηση των δεδομένων
2. Δημιουργία των γραμμών $L(\theta=0)$ και $L(\theta=2)$ και των δεδομένων για την εκπαίδευση του Νευρωνικού Δικτύου
 - a. Είσοδοι:
SES(k+1) ή Naive(k+1) για $L(\theta=2)$, περιορισμένη σε k σημεία
LRL(k+1) για την χρονοσειρά περιορισμένη σε k σημεία
 - b. Στόχος:
Data(k+1)
3. Δημιουργία των συνόλων εκπαίδευσης, δοκιμής και αξιολόγησης του νευρωνικού δικτύου.
4. Εκπαίδευση του Νευρωνικού Δικτύου
5. Επέκταση των γραμμών για τις πλήρεις γραμμές $L(\theta=0)$ και $L(\theta=2)$.
6. Παραγωγή προβλέψεων από το Νευρωνικό Δίκτυο με χρήση των τιμών του προηγούμενου βήματος ως είσοδο.
7. Εποχικοποίηση των προβλέψεων σύμφωνα με τους δείκτες που προσδιορίστηκαν στο Βήμα 1.

Πιο αναλυτικά:

1. Αρχικά διερευνάται αν η χρονοσειρά έχει εποχιακή συνιστώσα. Αν διαπιστωθεί πως έχει, η χρονοσειρά αποεποχικοποιείται με χρήση της κλασικής αποσύνθεσης (κάνοντας χρήση του πολλαπλασιαστικού μοντέλου).
2. Το νευρωνικό δίκτυο που θα καθορίσει τους συντελεστές των γραμμών $L(\theta=0)$ και $L(\theta=2)$ πρέπει να εκπαιδευτεί κατάλληλα. Για την εκπαίδευση του, η χρονοσειρά περιορίζεται στα k πρώτα στοιχεία της και χρησιμοποιούνται οι τιμές πρόβλεψης των $L(\theta=0)$ και $L(\theta=2)$ για το $k+1$ σημείο και ως στόχος ορίζεται το $k+1$ σημείο της χρονοσειράς.
3. Από τα δεδομένα του προηγούμενου σταδίου καθορίζονται τα σύνολα εκπαίδευσης, δοκιμής και αξιολόγησης του δικτύου. Εδώ τηρείται η χρονολογική σειρά και η αναλογία των συνόλων αυτών επί των συνολικών δεδομένων είναι 80 (εκπαίδευση) - 10 (δοκιμή) - 10 (αξιολόγηση) ή 70 (εκπαίδευση) - 15 (δοκιμή) - 15 (αξιολόγηση).
4. Με τα σύνολα που δημιουργήθηκαν πραγματοποιείται η εκπαίδευση του νευρωνικού δικτύου.
5. Για τις $L(\theta=0)$ και $L(\theta=2)$ γραμμές, παράγονται οι προβλέψεις σύμφωνα με την μέθοδο που έχουμε επιλέξει.
6. Τα αποτελέσματα του σταδίου 5 εισάγονται ως είσοδοι στο νευρωνικό δίκτυο.
7. Οι προβλέψεις που δίνει το νευρωνικό δίκτυο εποχικοποιούνται, αν χρειάζεται, σύμφωνα με το στάδιο 1.

Η εφαρμογή της μεθόδου Θ AI στον διαγωνισμό NN3.

Η μέθοδος ThetaAI δοκιμάστηκε πρώτη φορά στο πλαίσιο του διαγωνισμού NN3. Τα αποτελέσματα ήταν ενθαρρυντικά, μιας και κατατάχθηκε πρώτη στον διαγωνισμό των ii χρονοσειρών. Ακολουθούν αναλυτικά τα αποτελέσματα που δόθηκαν από τους διοργανωτές, για τις μεθόδους Theta και ThetaAI. Στον πρώτο πίνακα αναλύονται αποτελέσματα για το πλήρες σύνολο χρονοσειρών του διαγωνισμού για τις iii χρονοσειρές (Πίνακας 5), ενώ στον δεύτερο τα αποτελέσματα στο περιορισμένο σύνολο των ii χρονοσειρών (Πίνακας 6).

Μετρική Ακρίβεια	Μέθοδος	Όλα τα Δεδομένα	Βραχεία Εποχιακή	Βραχεία Μη Εποχιακή	Μακρά Εποχιακή	Μακρά Μη Εποχιακή
sMAPE	Theta	14.89%	14.25%	18.89%	9.68%	17.24%
	Theta AI	15.66%	14.27%	22.59%	9.53%	17.39%
	Theta AI	8.60%	9.13%	14.52%	5.31%	8.28%

Πίνακας 2: 111 Χρονοσειρές

Μετρική Ακρίβειας	Μέθοδος	Όλα τα Δεδομένα
sMAPE	Theta	13.70%
	Theta AI	13.07%

Πίνακας 3: 11 Χρονοσειρές

Ενώ τα αποτελέσματα της μεθόδου για τις 11 χρονοσειρές υπήρξαν πολύ καλά, ήταν λιγότερο ικανοποιητικά για το πλήρες σύνολο. Ο κύριος λόγος της διαφοράς στην απόδοση των δύο μεθόδων ήταν η χρήση διαφορετικών εργαλείων για την παραγωγή των προβλέψεων στη μια και στην άλλη περίπτωση. Για την παραγωγή των προβλέψεων στο σύνολο των 11 χρονοσειρών χρησιμοποιήθηκε το σύστημα **Πυθία** για την αποεποχικοποίηση των χρονοσειρών και το λογισμικό **Alyuda NeuroIntelligence** που παρέχει την δυνατότητα δημιουργίας νευρωνικών δικτύων για πρόβλεψη και κατηγοριοποίηση. Επειδή η παραπάνω διαδικασία ήταν μη αυτοματοποιημένη, για την παραγωγή των προβλέψεων των 111 χρονοσειρών χρησιμοποιήθηκε ειδικό λογισμικό που συνδύαζε τις βιβλιοθήκες του συστήματος **Πυθία** και την βιβλιοθήκη **Alyuda Neurofusion**. Η αυτοματοποιημένη μέθοδος δεν απέδωσε τόσο καλά όσο η διαδικασία που εφαρμόστηκε στις 11 χρονοσειρές.

6. Βελτίωση των αποτελεσμάτων του διαγωνισμού NN3

Η καλή απόδοση της μη αυτοματοποιημένης μεθόδου ThetaAI στο σύνολο των 11 χρονοσειρών του διαγωνισμού NN3 αποτέλεσε το κίνητρο για την εφαρμογή της στο διευρυμένο σύνολο των 111 χρονοσειρών του διαγωνισμού NN3. Στα πλαίσια αυτής της διπλωματικής εργασίας εφαρμόστηκε στο διευρυμένο σύνολο χρονοσειρών του NN3 η μη αυτοματοποιημένη μέθοδος ThetaAI με χρήση του λογισμικού *Alyuda NeuroIntelligence*, για να ελέγξουμε αν όντως θα έχει βελτιωμένη απόδοση σε σχέση με την κλασική μέθοδο θ. Επίσης, εφαρμόστηκε η αυτοματοποιημένη μέθοδος στα μηνιαία δεδομένα του διαγωνισμού M3 για να ελέγξουμε την απόδοση της σε αυτά. Τα αποτελέσματα αυτών των δοκιμών φαίνονται στον Πίνακα 7.

Μετρική Ακριβείας	Μέθοδος	Δεδομένα NN3 (111 χρονοσειρές - ορίζοντας πρόβλεψης 12)	Μηνιαία Δεδομένα M3 (1428 χρονοσειρές - ορίζοντας πρόβλεψης 18)
sMAPE	Theta	14.89%	13.88%
	Theta AI	14.53%	15.72%

Πίνακας 4

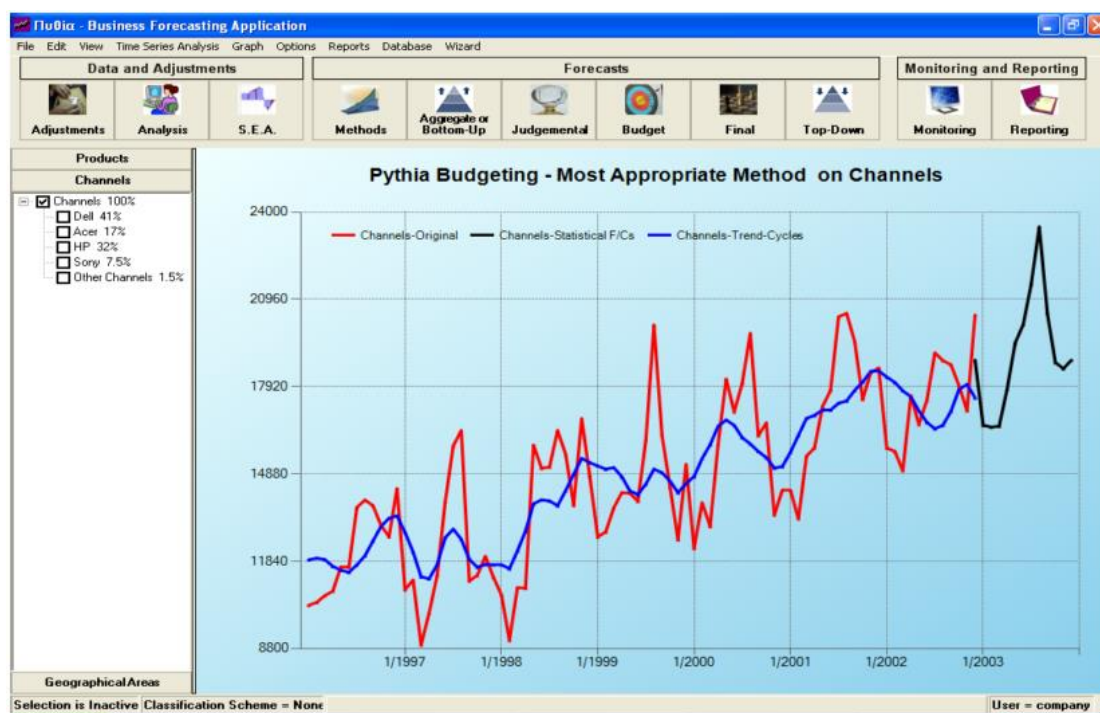
Παρατηρούμε πως η μέθοδος ThetaAI με χρήση της μη αυτοματοποιημένης μεθόδου, βελτίωσε την απόδοση της κλασικής Μεθόδου Θ για το σύνολο των 111 χρονοσειρών του διαγωνισμού NN3, όπως είχε συμβεί και στον NN3 με στο σύνολο των 11 χρονοσειρών. Για τα μηνιαία δεδομένα του διαγωνισμού M3, η ThetaAI, με αυτοματοποιημένη διαδικασία δεν απέδωσε καλύτερα από την κλασική μέθοδο θ. Το γεγονός που οφείλεται στον τρόπο εκπαίδευσης του νευρωνικού δικτύου από την βιβλιοθήκη Alyuda Neurosfusion. Τα σύνολα εκπαίδευσης, δοκιμής και αξιολόγησης του δικτύου δημιουργούνται αυτόματα και όχι τηρώντας χρονολογική σειρά όπως είναι επιθυμητό και όπως γίνεται στην μη αυτοματοποιημένη μέθοδο με χρήση του λογισμικού Alyuda NeuroIntelligence.

Ακολουθεί η αναλυτική περιγραφή των εργαλείων που χρησιμοποιήθηκαν.

ΕΡΓΑΛΕΙΑ

ΣΥΣΤΗΜΑ ΠΑΡΑΓΩΓΗΣ ΠΡΟΒΛΕΨΕΩΝ ΠΥΘΙΑ

Για την παραγωγή των βοηθητικών προβλέψεων των σταδίων 2 και 5 χρησιμοποιήθηκε η βιβλιοθήκη του συστήματος παραγωγής προβλέψεων Πυθία που έχει αναπτυχθεί από την Μονάδα Προβλέψεων και Στρατηγικής της Σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Εθνικού Μετσόβιου Πολυτεχνείου. Έτσι η εφαρμογή των μεθόδων SES, των απλών γραμμικών παλινδρομήσεων και οι αποεποχικοποιήσεις και εποχικοποιήσεις έγιναν με τις βιβλιοθήκες του συστήματος Πυθία.



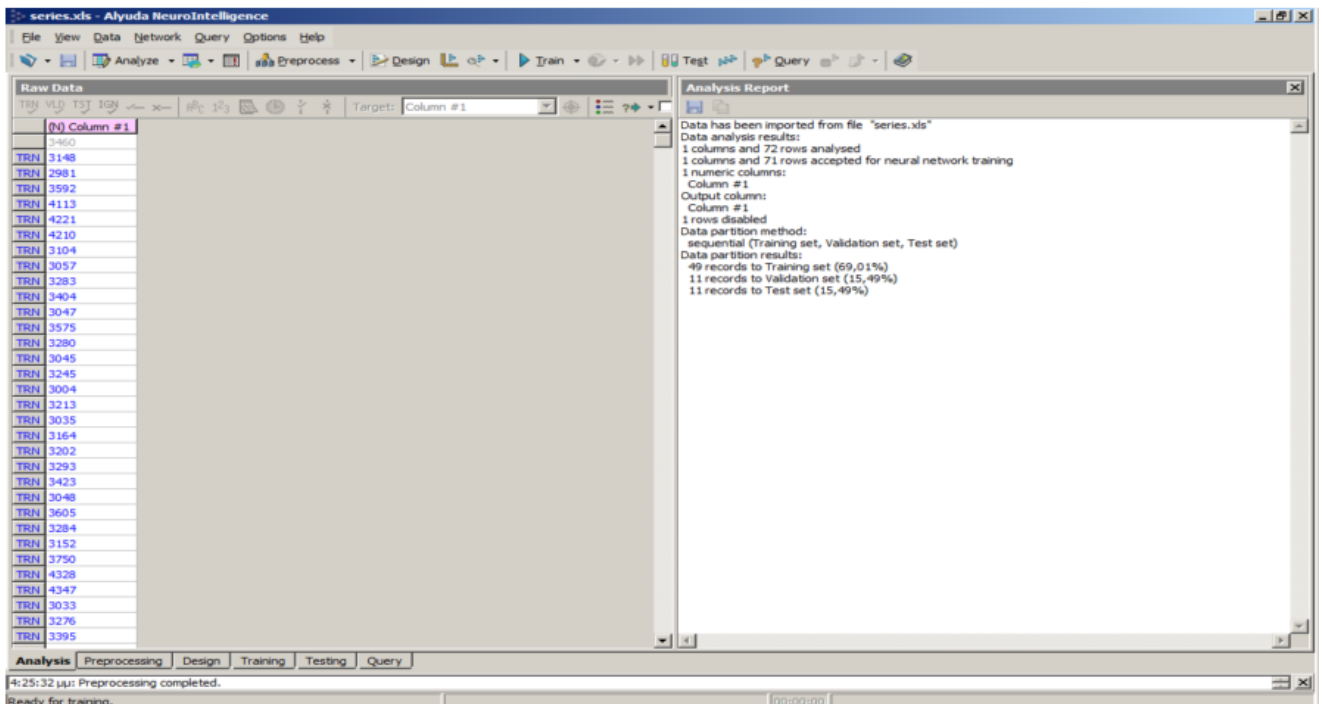
Το Alyuda NeuroIntelligence είναι λογισμικό το οποίο επιτρέπει την χρήση νευρωνικών δικτύων για πρόβλεψη και κατηγοριοποίηση. Το λογισμικό παρέχει πολλές δυνατότητες και επιδέχεται παραμετροποίηση. Οι εισοδοι και τα σύνολα εκπαίδευσης, δοκιμής και ελέγχου προήλθαν από το σύστημα Πυθία. Για κάθε μια από της 111 χρονοσειρές του πλήρους συνόλου, ακολουθήθηκαν τα βήματα της μεθόδου. Στο Πυθία έγιναν οι αποεποχικοποιήσεις και οι προβλέψεις για τις χρονοσειρές (τις περιορισμένες μέχρι το σημείο k αλλά τις ολόκληρες). Στη συνέχεια δημιουργήθηκε ένα αρχείο excel που διάβαζε το Alyuda NeuroIntelligence για την δημιουργία του νευρωνικού δικτύου.

Η διαδικασία που ακολουθείται είναι η εξής:

Στάδιο 1: Ανάλυση – Κατάτμηση (Analyze – Partition)

- Αρχικά επιλέγεται μια στήλη ως "target" και τίθεται ως η αρχική στήλη. Επιλέγεται αυτή που το Νευρωνικό Δίκτυο θεωρεί πως η μελέτη των δεδομένων της θα φέρει καλύτερη προσομοίωση στο πρόγραμμα. Μπορούμε να κάνουμε και αυθαίρετη επιλογή ώστε να προβούμε σε δοκιμές.
- Αναγνωρίζονται από το πρόγραμμα τα δεδομένα που αποτελούν **ασυνήθιστες τιμές (outliers)** και απομονώνονται από το υπόλοιπο της διαδικασίας. Η επιλογή outliers μπορεί να γίνει και από τον χρήστη, κατά την κρίση του.
- Αν πρόκειται για **χρονοσειρά** τότε υπάρχει ειδική ένδειξη επιλογής.
- Τα δεδομένα χωρίζονται σε **3 σύνολα**: σύνολο εκπαίδευσης, σύνολο δοκιμής και σύνολο ελέγχου. Αν πρόκειται για χρονοσειρά, *διατηρείται μια κατάλληλη χρονολογική δομή*.

- Στις επιλογές της "Analyze" (ανάλυσης) μπορεί να γίνει η επιλογή είτε το πρόγραμμα να αγνοεί είτε να αντικαθιστά τα πιθανά κενά στα δεδομένα και τις ακραίες τιμές (outliers). Συγχρόνως στις επιλογές της "Partition" (κατάτμησης) καθορίζεται το ποσοστό που κατέχει το κάθε σετ επί του συνόλου (το **σετ εκπαίδευσης** 70%, το **σετ δοκιμής** 15% και το **σετ αξιολόγησης** 15%).



Στάδιο 2: Επανελέγχος

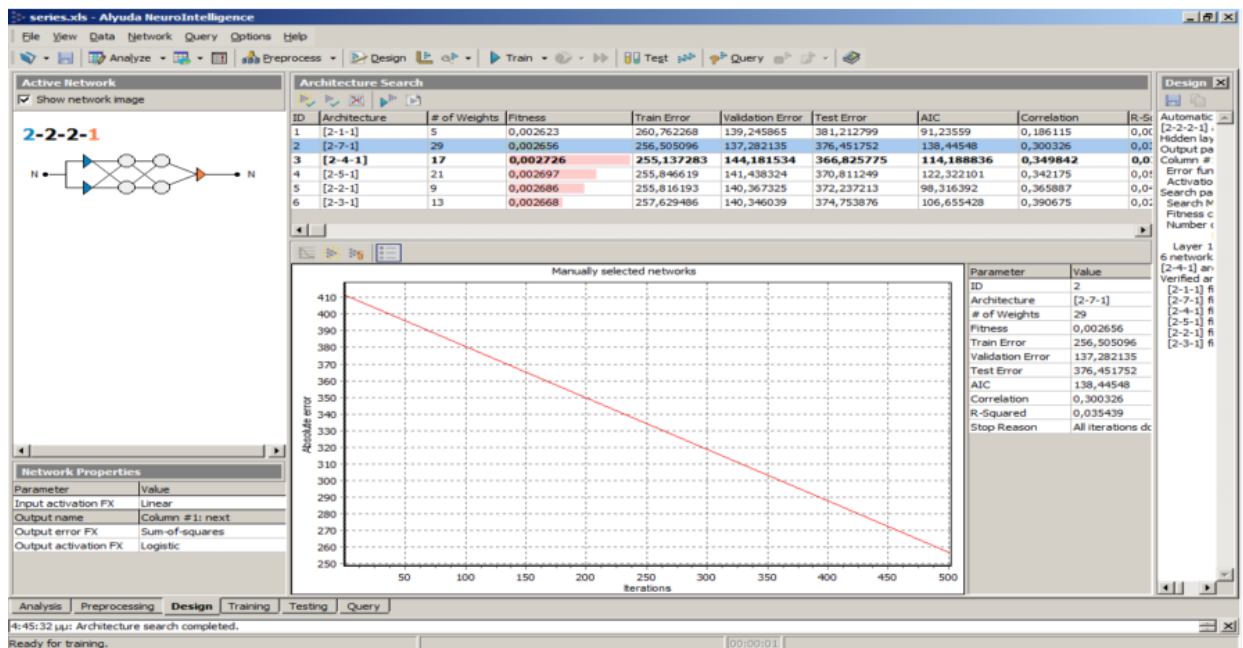
Το πρόγραμμα κατηγοριοποιεί τα δεδομένα και εμφανίζει χαρακτηριστικά στοιχεία κάθε στήλης ξεχωριστά, όπως ο Μέσος Όρος, η Μέγιστη και Ελάχιστη Τιμή, το χαρακτηρισμό της στήλης ως Input ή Output (δεδομένα).

Με την επιλογή Options μπορεί να γίνει οποιαδήποτε ρύθμιση ώστε να αποκλείονται τιμές που παραβιάζουν κάποιο κριτήριο. Ενδεικτικά, μπορούν να αποκλειστούν τιμές που υπερβαίνουν ένα ορισμένο όριο που τίθεται από τον χρήστη (π.χ απόκλιση από τη μέγιστη τιμή μέχρι 3 μονάδες).

Στάδιο 3: Διερεύνηση και σχεδιασμός της Αρχιτεκτονικής

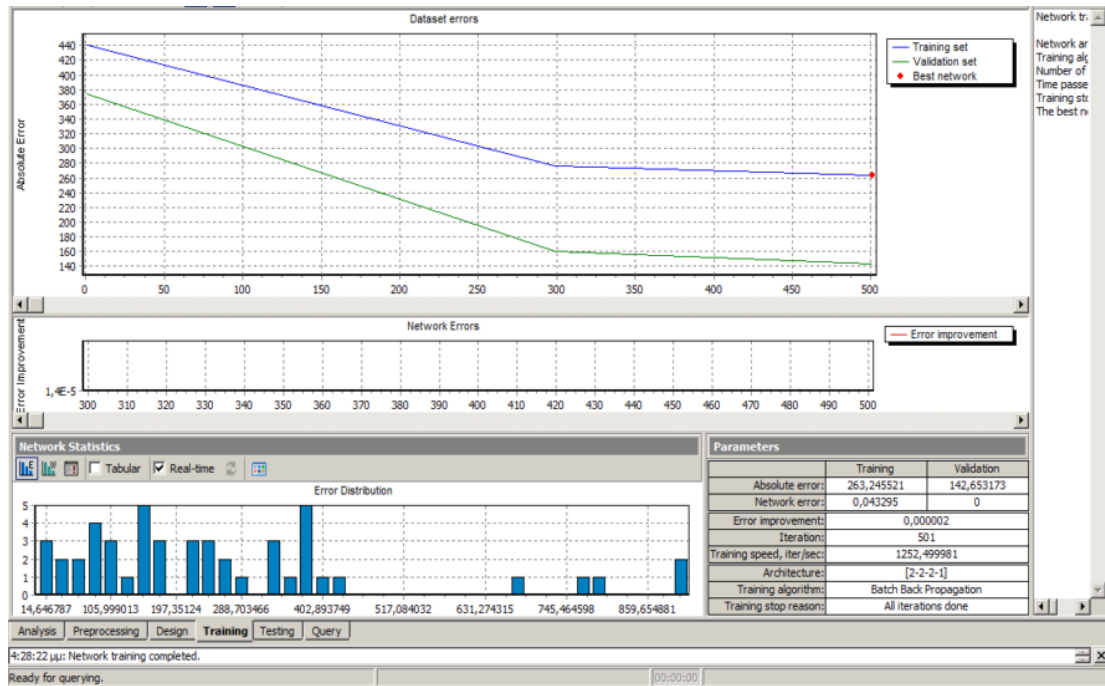
- Το στάδιο αυτό αποτελεί ίσως το **σημαντικότερο** όλου του προγράμματος. Αρχικά επιλέγεται αν θα εφαρμοστεί **ευριστική** ή **αναλυτική** μέθοδος.

- Στο στάδιο αυτό καθορίζονται παράμετροι όπως ο **αριθμός των κρυφών στρωμάτων**, των **μονάδων** κάθε στρώματος, ο αριθμός των **επαναλήψεων** καθώς και το **πόσες φορές** εξετάζεται η **μεθοδολογία** σε κάθε επανάληψη.
- **Επιλέγεται η καλύτερη μέθοδος** σύμφωνα με τις επιλογές που κάναμε και σύμφωνα με ένα **προκαθορισμένο κριτήριο** που επιλέγεται από τον χρήστη (Fitness, correlation, R-squared).



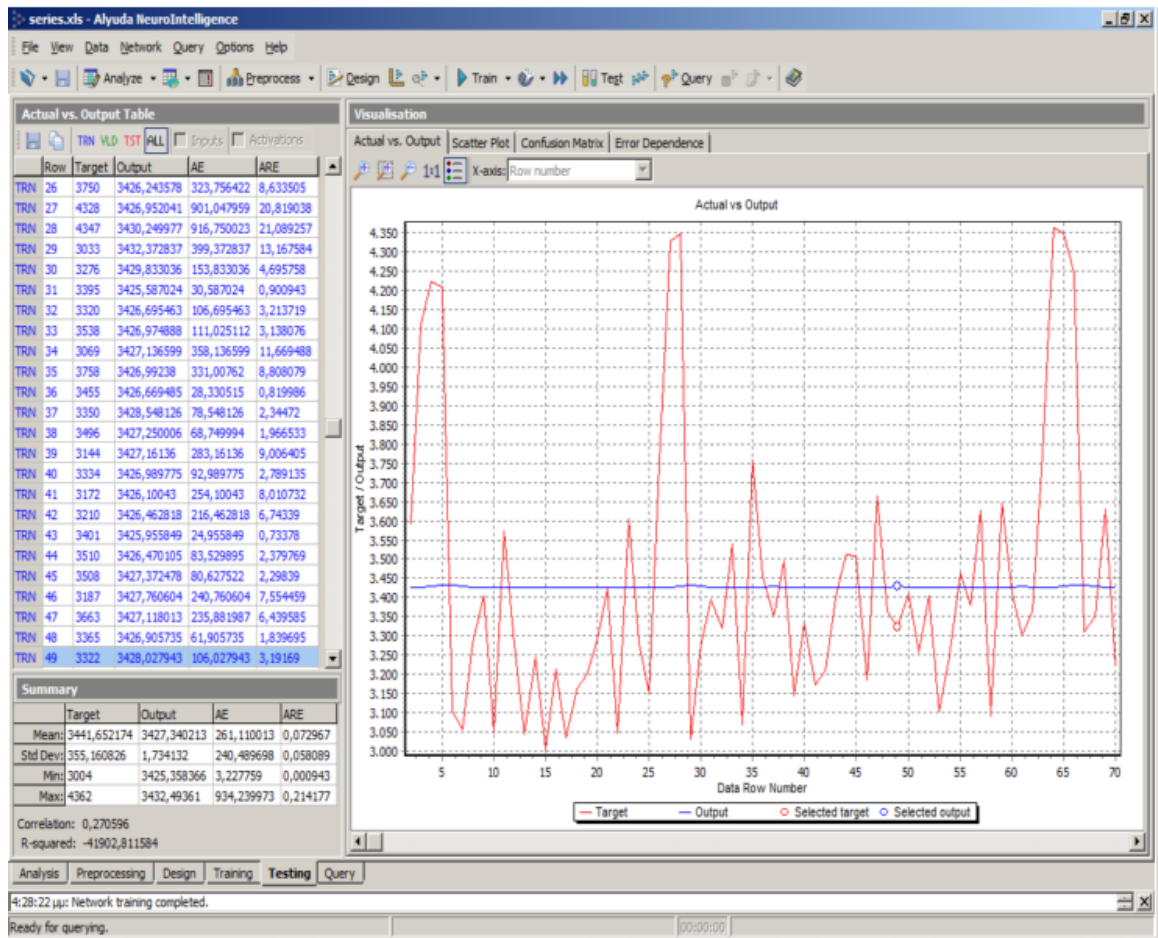
- Το πρόγραμμα εμφανίζει τα **αποτελέσματα** του προηγούμενου σταδίου, οπότε προχωράμε στην **ενεργοποίηση** της επιλεγμένης αρχιτεκτονικής. Σε περίπτωση που η αρχιτεκτονική δε μας ικανοποιεί, προχωρούμε **στην επανασχεδίαση** της, αλλάζοντας την τιμή κάποιας παραμέτρου.
- Επισημαίνεται ότι αντί της ανωτέρω διαδικασίας, υπάρχει η δυνατότητα της **αριστοποίησης** μιας συγκεκριμένης αρχιτεκτονικής, σύμφωνα με τις γνώσεις μας και την ευριστική μας προσέγγιση.
- Στη φάση αυτή εμφανίζεται ένα ιδιαίτερα ενδιαφέρον **Γράφημα** που ουσιαστικά δείχνει τη **δομή** και την **αρχιτεκτονική** του **νευρωνικού δικτύου** που επιλέγουμε με τον έναν ή τον άλλον τρόπο.

Στάδιο 4: Εκπαίδευση



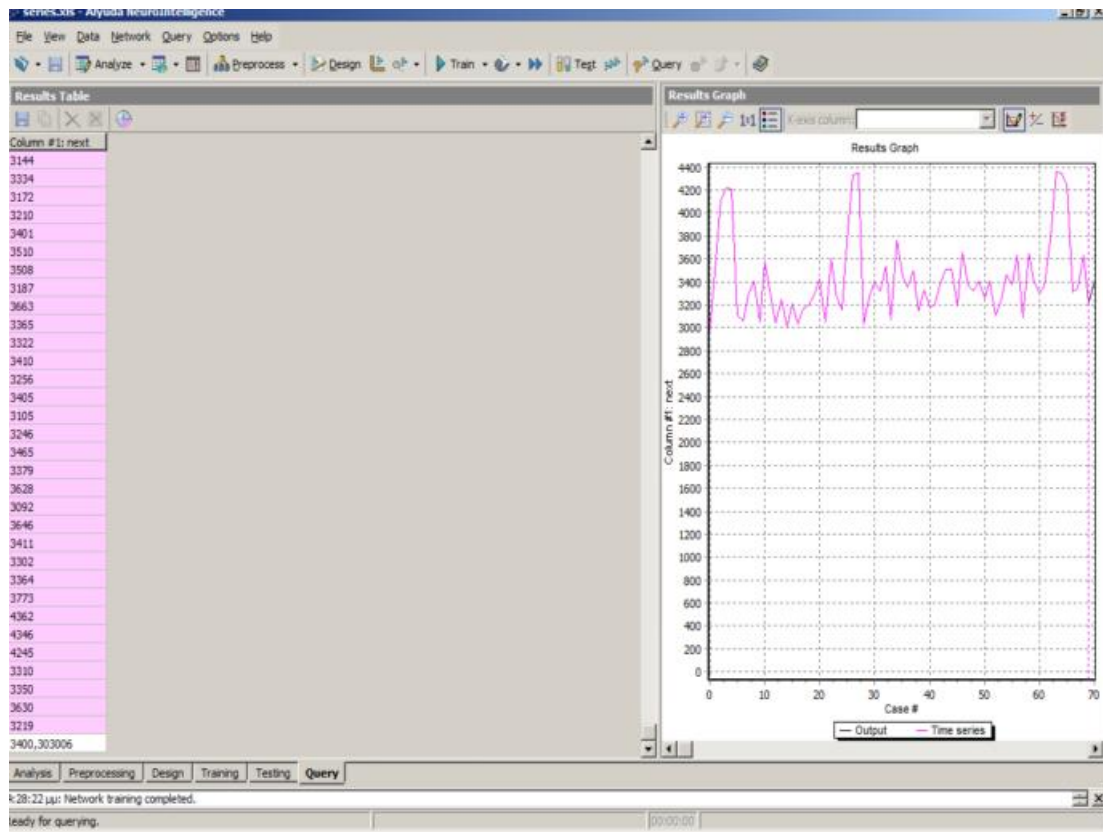
Στάδιο 5: Αξιολόγηση

- Κατά το στάδιο αυτό, αρχικά εμφανίζεται το πιο σημαντικό **Γράφημα** όσον αφορά τα **αποτελέσματα** της έρευνας μας για την κατάλληλη επιλογή αρχιτεκτονικής στο δίκτυο (με διαφορετικό χρώμα οι αρχικές και οι τελικές προβλέψεις μας, "actual-output").



- Εμφανίζονται σε Πίνακα, **τιμές σημαντικών παραμέτρων** των δεδομένων του συνόλου ελέγχου, όπως το **ΑΕ** ή το **ΑΡΕ**.
- Αν δεν είμαστε ικανοποιημένοι από το συνολικό αποτέλεσμα, ή μας εμφανίζονται μεγάλες αποκλίσεις στα επιθυμητά μας αποτελέσματα, μπορούμε να **επαναλάβουμε** κάποια από τα προηγούμενα **Βήματα - Στάδια**, κάτι που αποτελεί **συχνό φαινόμενο** και είναι ένας από τους παράγοντες (όπως και η συλλογή δεδομένων), που η **επιλογή** του κατάλληλου **Νευρωνικού Δικτύου**, αποτελεί μια **διαδικασία** ιδιαίτερα **χρονοβόρα** και **επίπονη**.

Στάδιο 6: Εξαγωγή Αποτελεσμάτων



Το Alyuda NeuroFusion είναι μια βιβλιοθήκη ρουτινών .NET που επιτρέπουν την δημιουργία, εκπαίδευση και αποτίμηση νευρωνικών δικτύων μέσω κώδικα. Χρησιμοποιήθηκε για την παραγωγή των προβλέψεων των μηνιαίων χρονοσειρών του διαγωνισμού M3. Με την βοήθεια λογισμικού που χρησιμοποιούσε το σύστημα Πυθία για αποεχικοποίηση και εποχικοποίηση των δεδομένων και χρήση της βιβλιοθήκης Alyuda NeuroFusion για το κομμάτι των ΝΔ μπορούν να παράγονται αυτοματοποιημένα αποτελέσματα για την ThetaAI. Όπως προαναφέρθηκε, το βασικό μειονεκτήματα της βιβλιοθήκης ως προς την χρήση της σε νευρωνικά δίκτυα για χρονοσειρές είναι ότι επιλέγει τυχαία τα σύνολα εκπαίδευσης, δοκιμής και αξιολόγησης πράγμα που στην περίπτωση των χρονοσειρών δεν είναι επιθυμητό.

Τα τελευταία δεδομένα της χρονοσειράς είναι μεγαλύτερης βαρύτητας και είναι καλό να χρησιμοποιούνται για την δοκιμή και την αξιολόγηση.

7. ΣΥΜΠΕΡΑΣΜΑΤΑ – ΜΕΛΛΟΝΤΙΚΕΣ ΕΠΕΚΤΑΣΕΙΣ

Από τα αποτελέσματα της εργασίας αλλά και από την διαδικασία που χρειάστηκε να ακολουθηθεί βγάζουμε τα ακόλουθα συμπεράσματα:

Η χρήση τεχνικών και εργαλείων Τεχνητής Νοημοσύνης μπορεί να βελτιώσει την απόδοση των κλασικών μεθόδων πρόβλεψης

Η χρήση τεχνικών και εργαλείων Τεχνητής Νοημοσύνης είναι χρονοβόρα διαδικασία και ως εκ τούτου μειονεκτεί σε σχέση με τις κλασικές μεθόδους πρόβλεψης όταν ζητούμενο είναι η μαζική παραγωγή προβλέψεων.

Μειονέκτημα είναι ότι η μέθοδος δεν αναπαράγει ακριβώς τα ίδια αποτελέσματα επειδή στηρίζεται σε νευρωνικά δίκτυα.

Για βελτίωση και καλύτερη τεκμηρίωση των αποτελεσμάτων της ThetaAI αλλά και συναφή έρευνα προτείνεται:

- Η επέκταση και βελτίωση της μεθόδου ThetaAI ώστε η πρόβλεψη να γίνεται βάσει περισσότερων του ενός σημείου και όχι μόνον του αμέσως προηγούμενου.
- Η χρήση περισσότερων γραμμών, εκτός της γραμμή "0" και "2" κατά την υλοποίηση της μεθόδου ThetaAI.
- Η χρήση μεγάλων συνόλων χρονοσειρών για τη διεξαγωγή πειραμάτων, και με τρόπο που να λαμβάνεται υπ' όψιν η χρονολογική δομή των δεδομένων κατά τη διαμόρφωση του νευρωνικού δικτύου.
- Η χρήση διαφορετικών εργαλείων αντί για νευρωνικά δίκτυα. Ενδεικτικά, μπορούν να χρησιμοποιηθούν τεχνικής ασαφούς λογικής (fuzzy logic).

Βιβλιογραφία

- *Armstrong S., Adya M., Collopy F. 2001. Principles of Forecasting NN: Rule-based forecasting: Using Judgement in time-series extrapolation*
- *Assimakopoulos V., Nikolopoulos K., 2000. The Theta model: a decomposition approach to forecasting International Journal of Forecasting, 16, 521–530.*
- *Crone S., 2005. Forecasting with Artificial Neural Networks EVIC 2005 Tutorial Santiago de Chile.*
- *Makridakis, S., 1986. The art and science of forecasting: An assessment and future directions, International Journal of Forecasting, 2, 15-39.*
- *Makridakis, S., 1993. Accuracy measures: theoretical and practical concerns, International Journal of Forecasting, 9, 527–529.*
- *Makridakis S., Wheelwright S., Hyndman R., 1998. Forecasting: Methods and Applications 3rd edition, John Wiley & Sons, New York*
- *Makridakis S., Hibon M., 2000. The M3-Competition: Results, conclusions and implications. International Journal of Forecasting, 164, 451– 476.*
- *Makridakis S., Hogarth R., Gaba A., 2008. Dance with Chance, Making Luck work for You, Oneworld Publications,*
- *Nikolopoulos, K., 2002. Business Forecasting Methodology and Information System. Thesis (PhD). National Technical University of Athens.*
- *Ασημακόπουλος, Β., 2005. Μέθοδοι Προβλέψεων. Αθήνα: Πανεπιστημιακές Εκδόσεις ΕΜΠ.*
- *Δημόπουλος, Σ., 2008. Πρόβλεψη χρονοσειρών με τη χρήση νευρωνικών δικτύων-εφαρμογή στον τομέα του τουρισμού*
- *Πετρόπουλος, Φ., 2007. Βελτιστοποίηση μεθόδου πρόβλεψης Θ και εφαρμογή επί των μηνιαίων δεδομένων των διαγωνισμών M3 και T.*
- *Τσιάφα Ε., 2008. Βελτιστοποίηση Μεθόδου Πρόβλεψης Θ με Χρήση Νευρωνικών Δικτύων*