



**ΠΟΛΥΤΕΧΝΕΙΟ
ΚΡΗΤΗΣ**

ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΑΡΑΓΩΓΗΣ & ΔΙΟΙΚΗΣΗΣ

ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ ΤΩΝ ΜΕΘΟΔΩΝ ΠΡΟΒΛΕΨΗΣ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ ΥΠΟ

ΡΑΛΛΗ ΙΩΑΝΝΗ

Π ρ ό λ ο γ ο ς

Η παρούσα διπλωματική εργασία εκπονήθηκε στο Τμήμα Μηχανικών Παραγωγής και Διοίκησης του Πολυτεχνείου Κρήτης. Η εκκίνηση της διπλωματικής εργασίας τοποθετείται χρονικά τον Φεβρουάριο του 2014.

Στο σημείο αυτό θα ήθελα να ευχαριστήσω τον καθηγητή μου Γ. Ατσαλάκη για την καθοδήγησή και την στήριξη του κατά την εκπόνηση της παρούσας διπλωματικής εργασίας.

Ιούνιος 2014

Ράλλης Ιωάννης

ΠΕΡΙΕΧΟΜΕΝΑ

| | |
|--|-----------|
| Πρόβλεψη..... | 7 |
| Ποιοτικές Μέθοδοι..... | 12 |
| Μέθοδος Delphi..... | 13 |
| Εφαρμογή της Delphi..... | 16 |
| Πλεονεκτήματα και περιορισμοί | 20 |
| Δομημένη Αναλογία | 22 |
| Αιτιοκρατικό Μοντέλο | 24 |
| Μέθοδοι Πρόβλεψης Συνεχούς Ζήτησης | 25 |
| Μοντέλο Χρονοσειρών | 25 |
| Ποιοτικά χαρακτηριστικά χρονοσειρών – | |
| Στατιστικοί δείκτες & δείκτες σφαλμάτων | 29 |
| Κλασσική μέθοδος Αποσύνθεσης..... | 33 |
| Κριτήρια αξιολόγησης των μεθόδων πρόβλεψης | 38 |
| Μέθοδοι εξομάλυνσης | 39 |
| Στατιστικές μέθοδοι Πρόβλεψης..... | 40 |
| Μέθοδος Naïve..... | 40 |
| Μέθοδοι εκθετικής εξομάλυνσης..... | 41 |
| Απλή εκθετική εξομάλυνση σταθερού επιπέδου (SES)..... | 41 |
| Εκθετική εξομάλυνση γραμμικής τάσης..... | 43 |
| Εκθετική εξομάλυνση μη γραμμικής τάσης..... | 45 |
| Μοντέλα παλινδρόμησης (Regression Models) | 47 |
| Απλή γραμμική παλινδρόμηση | 48 |
| Συντελεστής γραμμικής συσχέτισης..... | 51 |

| | |
|---|-----------|
| Μοντέλο πολλαπλής γραμμικής παλινδρόμησης..... | 52 |
| Μοντέλο Theta..... | 53 |
| Μοντέλα ARIMA..... | 55 |
| Βασικές έννοιες προχωρημένης ανάλυσης..... | 56 |
| Η συνάρτηση αυτοσυσχέτισης (autocorrelation function) (ACF)..... | 56 |
| Μοντέλο λευκού θορύβου | 57 |
| Δειγματική κατανομή των αυτοσυσχετίσεων | 57 |
| Στατιστικοί Δείκτες Q, Q^* | 58 |
| Συντελεστές μερικής αυτοσυσχέτισης..... | 59 |
| Έλεγχος στασιμότητας της χρονοσειράς..... | 59 |
| Μετατροπή μη στάσιμης χρονοσειράς σε στάσιμη | 60 |
| Μοντέλο τυχαίου περιπάτου..... | 61 |
| Εποχιακή Διαφόριση..... | 61 |
| Μοντέλα χρονοσειρών ARIMA | 62 |
| Αυτοπαλινδρομικά μοντέλα πρώτης τάξης | 64 |
| Μοντέλα κινητού μέσου όρου πρώτης τάξης..... | 65 |
| Διαδικασίες ARMA | 65 |
| Διαδικασίες ARIMA..... | 66 |
| Εποχικότητα και μοντέλα ARIMA..... | 66 |
| Εκτίμηση παραμέτρων | 67 |
| Επιλογή βέλτιστου μοντέλου ARIMA | 68 |
| Διαγνωστικός έλεγχος | 69 |
| Διαφόριση και πρόβλεψη..... | 70 |
| Μέθοδοι Διακοπτόμενης Ζήτησης..... | 72 |

| | |
|---|------------|
| Μέθοδος CROSTON | 72 |
| Μέθοδος Syntetos & Boylan Approximation (SBA) | 74 |
| Επιλογή κατάλληλης μεθόδου πρόβλεψης | 75 |
| Νευρωνικά Δίκτυα | 77 |
| Πλεονεκτήματα νευρωνικών δικτύων | 77 |
| Τεχνητή Νοημοσύνη | 81 |
| Περιοχές Έρευνας Τεχνητής Νοημοσύνης | 83 |
| Έμπειρα Συστήματα | 83 |
| Γενετικοί Αλγόριθμοι | 85 |
| Ασαφής Λογική | 86 |
| Ιδιότητες Ασαφών συνόλων | 87 |
| Λογικές πράξεις στα Ασαφή σύνολα | 94 |
| Ασαφείς σχέσεις | 96 |
| Ασαφής συλλογιστική | 97 |
| Συστήματα ασαφούς συλλογιστικής-νευροασαφή συστήματα | 100 |
| Συστήματα τύπου Mamdani | 105 |
| Συστήματα τύπου Sugeno | 108 |
| Νευροασαφή συστήματα | 111 |
| Ασαφείς νευρώνες | 111 |
| Νευρώνες συμμετοχής | 111 |
| Λειτουργικοί νευρώνες | 112 |
| Συνθετικοί νευρώνες | 113 |
| Προσαρμοστικά Δίκτυα (Adaptive networks) | 113 |
| Αρχιτεκτονική του ANFIS | 114 |

| | |
|---|------------|
| Τεχνητά Νευρωνικά Δίκτυα..... | 117 |
| Ιστορική εξέλιξη νευρωνικών δικτύων..... | 122 |
| Εκπαίδευση νευρωνικού Δικτύου | 124 |
| Μέθοδος ελαχίστου μέσου τετραγωνικού λάθους..... | 125 |
| Μέθοδος ταχύτερης καθόδου..... | 127 |
| Αλγόριθμος οπισθοδρόμησης..... | 127 |
| Αρχιτεκτονικές ΤΝΔ..... | 128 |
| Νευρωνικό οπίσθιας τροφοδότησης | 129 |
| Νευρωνικό δίκτυο εμπρόσθιας τροφοδότησης | 129 |
| Νευρωνικό δίκτυο παλινδρόμησης | 129 |
| Πιθανοκρατικό Νευρωνικό δίκτυο | 129 |
| Νευρωνικό δίκτυο ακτινικής συνάρτησης | 129 |
| Μηχανές διανυσμάτων Υποστήριξης βάσης..... | 130 |
| Μετάδοση σήματος στο νευρωνικό δίκτυο..... | 132 |
| Πλεονεκτήματα και μειονεκτήματα ΤΝΔ | 138 |
| Συμπεράσματα | 141 |
| Βιβλιογραφία..... | 143 |

ΠΡΟΒΛΕΨΗ

Αντικείμενο της επιστήμης των προβλέψεων είναι η συστηματοποίηση της παραγωγής και η ερμηνεία της αλλαγής του περιβάλλοντος (οικονομικό, κυτταρικό κλπ). Μέσω της αξιοποίησης της διαθέσιμης γνώσης για την παραγωγή ισχυρισμών και την επεξεργασία της πληροφορίας προσπαθούμε να προβλέψουμε μεγέθη, τάσεις, οικονομικούς δείκτες. Η πρόκληση που ανακύπτει, ο στοχαστικός χαρακτήρας του προβλήματος καθιστούν τον τομέα των προβλέψεων ιδιαίτερα δύσκολο αλλά συγχρόνως και ενδιαφέρον. Ανεξάρτητα από τις συνθήκες ή χρονικούς ορίζοντες που εμπλέκονται, οι προβλέψεις αποτελούν ένα σημαντικό βοήθημα για τον αποτελεσματικό και αποδοτικό σχεδιασμό. Η προβλεψιμότητα ενός γεγονότος ή μιας ποσότητας εξαρτάται από πολλούς παράγοντες, όπως οι εξής: το πόσο καλά καταλαβαίνουμε τους παράγοντες που συμβάλλουν στην μεταβολή, πόσο υπάρχουν διαθέσιμα στοιχεία-δεδομένα για την παραγωγή εκτιμήσεων και αν οι προβλέψεις μπορούν να επηρεάσουν το αντικείμενο πρόβλεψης.

Πολλοί άνθρωποι εσφαλμένα υποθέτουν ότι οι προβλέψεις δεν είναι δυνατόν να επαληθευτούν σε ένα μεταβαλλόμενο περιβάλλον. Κάθε περιβάλλον αλλάζει, και ένα καλό μοντέλο πρόβλεψης αποτυπώνει τον τρόπο με τον οποίο τα πράγματα αλλάζουν. Οι προβλέψεις σπάνια υποθέτουν ότι το περιβάλλον είναι αμετάβλητο. Αυτό που υποθέτουμε είναι ότι ο τρόπος με τον οποίο το περιβάλλον αλλάζει. Δηλαδή, ένα εξαιρετικά ασταθές περιβάλλον θα συνεχίσει να είναι εξαιρετικά ασταθές, μια επιχείρηση με κυμαινόμενες πωλήσεις θα συνεχίσει να έχει διακυμάνσεις των πωλήσεων. Ένα μοντέλο πρόβλεψης έχει σκοπό να συλλάβει τον τρόπο που κινούνται οι εξελίξεις, είτε πρόκειται για μετοχές, είτε πρόκειται για τεχνολογικές εξελίξεις.

Οι προβλέψεις καταστάσεων διαφέρουν σε μεγάλο βαθμό στους χρονικούς ορίζοντες, στους παράγοντες που καθορίζουν τα πραγματικά αποτελέσματα, τα είδη των δεδομένων, και

πολλές άλλες πτυχές . Μέθοδοι πρόβλεψης μπορεί να είναι πολύ απλές, όπως χρησιμοποιώντας την πιο πρόσφατη παρατήρηση ως πρόβλεψη η οποία ονομάζεται « απλοϊκή μέθοδο » (Naïve method) , ή εξαιρετικά πολύπλοκη, όπως τα νευρωνικά δίκτυα τα νευρο-ασαφή συστήματα και τα οικονομετρικά συστήματα εξισώσεων.

Υπάρχουν πολλές κατηγορίες τεχνολογικής πρόβλεψης, οι οποίες κατηγοριοποιούνται σε 8 ομάδες :

Η γνώμη των ειδικών, Ανάλυση των τάσεων, Παρακολούθηση και νοημοσύνη , μοντελοποίηση και προσομοίωση, σενάρια, Στατιστική, Περιγραφική μέθοδος , Δημιουργικότητα και αποτίμηση / απόφαση / Οικονομικοί Μέθοδοι.

Γνώμη των ειδικών (Expert opinion) :

- Delphi (επαναληπτική μέθοδος)
- Ομάδες εστίασης (focus groups , εργαστήρια)
- συνεντεύξεις
- συμμετοχικές τεχνικές

Ανάλυση των τάσεων (Trend Analysis) :

- Trend Extrapolation (Growth Curve Fitting)
- Trend impact analysis (Ανάλυση των επιπτώσεων)
- Precursor Analysis (Ανάλυση προοδρόμου)
- Long Wave Analysis

Μέθοδοι Παρακολούθησης και Πληροφοριών :

- Παρακολούθηση (περιβαλλοντική ανίχνευση ,παρακολούθηση τεχνολογικών εξελίξεων)
- Βιβλιομετρία (εύρεση προφίλ , εύρεση των διπλωμάτων ευρεσιτεχνίας , εξόρυξη κειμένου)

Στατιστικές Μέθοδοι :

- Ανάλυση συσχέτισης (correlation analysis)
- Δημογραφικά
- Ανάλυση των επιπτώσεων σταυρού (cross impact analysis)
- Ανάλυση κινδύνου

Μοντελοποίηση και Προσομοίωση :

- Ανάλυση βιωσιμότητας (ανάλυση του κύκλου ζωής) Sustainability Analysis
- Αιτιώδης Μοντέλα (casual models)
- Μοντελοποίηση Διάχυσης (Diffusion modelling)
- Complex Adaptive System Modeling (CAS)
- Προσομοίωση Συστημάτων (System Dynamics)
- Τεχνολογική Αλλαγή (Technological Substitution)
- Σενάρια-προσομοίωσης (gaming , διαδραστικά σενάριο)

- Μοντελοποίηση οικονομικής βάσης (Economic base modeling)
- Αξιολόγηση των τεχνολογιών

Σενάρια με ελέγχους συνέπειας (διαχείριση σεναρίων) :

- Field Anomaly Relaxation Method (πεδίο μεθόδου ανώμαλης χαλάρωσης FAR)

Αποτίμηση / Απόφαση / Οικονομικά Μεγέθη :

- Δέντρα συνάφειας
- Ανάλυση Δράσης
- Ανάλυση Κόστους - Όφελους
- Ανάλυση Αποφάσεων (ανάλυση χρησιμότητας)
- Μοντελοποίηση οικονομικής βάσης (ανάλυση εισροών - εκροών)

Μέθοδοι Πινάκων και Περιγραφικοί (Descriptive and Matrices Methods)

- Αναλογίες
- Backcasting (ανάλυση εκ των υστέρων)
- Λίστα ελέγχου για την αναγνώριση των επιπτώσεων
- Μοντελοποίηση Καινοτομιών
- Θεσμική Ανάλυση

- Ανάλυση Μετριάσμων
- Μορφολογική Ανάλυση
- Roadmapping (προϊόντα τεχνολογίας roadmapping)
- Κοινωνικών επιπτώσεων
- Εκτίμηση πολλαπλών προοπτικών
- Οργανωτική Ανάλυση
- Ανάλυση απαιτήσεων (ανάλυση αναγκών)

Δημιουργικότητα :

- Brainstorming
- δημιουργικά εργαστήρια
- TRIZ (Kucharavy and De Guio, 2005)
- Όραμα γενίας (Vision Generation)
- Ανάλυση Επιστημονικής Φαντασίας

ΠΟΙΟΤΙΚΕΣ ΜΕΘΟΔΟΙ (Expert opinion)

Μια από τις απλούστερες και πιο διαδεδομένες ποιοτικές μεθόδους πρόβλεψης που εφαρμόζονται για τη λήψη επιχειρηματικών αποφάσεων είναι η μέθοδος της «γνώμης του συμβουλίου στελεχών». Η αρχή της μεθόδου είναι η πρόβλεψη να βασίζεται σε υποκειμενικές εκτιμήσεις που εκφράζουν κατάλληλα άτομα. Τα άτομα αυτά είναι στελέχη της επιχείρησης, που προέρχονται από όλα τα βασικά τμήματα (παραγωγή, πωλήσεις, χρηματοοικονομικό κλπ). Τα άτομα αυτά συνεδριάζουν και αποφασίζουν ομαδικά, με τη βοήθεια ποσοτικών στοιχείων, δεικτών κ.λπ., για το ποια είναι η καλύτερη πρόβλεψη. Πλεονέκτημα της μεθόδου είναι ότι οι προβλέψεις ετοιμάζονται γρήγορα και εύκολα, ότι βασίζονται στη γνώμη και στις πληροφορίες στελεχών με διαφορετική ειδικότητα και ρόλο και ότι συχνά είναι η μόνη εφικτή μέθοδος, πράγμα που συμβαίνει όταν δεν διατίθενται στατιστικά δεδομένα ή όταν η πρόβλεψη αφορά φαινόμενα σε ένα περιβάλλον που μεταβάλλεται ταχύτατα.

Μειονέκτημα της μεθόδου αυτής είναι το γεγονός ότι η βαρύτητα της γνώμης ενός στελέχους στο τελικό αποτέλεσμα είναι συνάρτηση της θέσης του στελέχους αυτού μέσα στην επιχείρηση και της προσωπικότητας του και όχι των πληροφοριών που διαθέτει ή της ικανότητας του να προβλέπει το μέλλον.

Μια παραλλαγή της μεθόδου είναι η περιοδική συγκέντρωση από το γενικό διευθυντή ή άλλο ανώτατο στέλεχος γραπτών εκτιμήσεων των στελεχών της επιχείρησης και η σύνθεση των εκτιμήσεων σε μία τελική αντιπροσωπευτική εκτίμηση.

Μέθοδος DELPHI

Η μέθοδος DELPHI εφαρμόζεται για τη διενέργεια τεχνολογικών προβλέψεων, δηλαδή μακροπρόθεσμων κυρίως προβλέψεων που αφορούν την τεχνολογία και το γενικό περιβάλλον, μέσα στο οποίο αναπτύσσεται η επιχειρηματική δράση.

Τα τελευταία τριάντα χρόνια γίνεται πολύς λόγος για τη μέθοδο DELPHI (Ανδρουλιδάκης, Σ. και Τσολάκης, Χρ.1990). Είναι μια μέθοδος-τεχνική που επινοήθηκε κατά τη δεκαετία του 1940 από τη RAND Corporation με την καθοδήγηση του Helmer (Dalkey, 1968). Η πρωτοποριακή μελέτη πραγματοποιήθηκε, για να επιτευχθεί μια σύγκλιση απόψεων μεταξύ διαφόρων ειδικών σχετικά με κάποιο θέμα άμυνας της Στρατιωτικής Αεροπορίας των Η.Π.Α. Φέρει το όνομα "DELPHI" από το Ελληνικό "ΔΕΛΦΟΙ", γιατί βοηθά στην πρόβλεψη. Στις αρχικές μελέτες γινόταν προσπάθεια με τη μέθοδο αυτή να δοθούν απαντήσεις σε ερωτήσεις για το μέλλον και σε περιπτώσεις που το θέμα ήταν περίπλοκο και αβέβαιο (Dalkey, 1968). Χρησιμοποιεί μια ομάδα ειδικών κατά τέτοιο τρόπο, ώστε να εξαλείφεται το ενδεχόμενο της κυριαρχίας αυτών με το μεγαλύτερο κύρος, πειθώ και δυναμισμό. Επιδιώκεται, δηλαδή, η εξουδετέρωση των φαινομένων που συνδέονται με την ομαδική συμπεριφορά και αποτελούν κίνδυνο, στον οποίον υπόκειται η μέθοδος της «γνώμης του συμβουλίου στελεχών». Το ζητούμενο είναι να επιτευχθεί τελικά η ομόφωνη γνώμη των ειδικών ως συγκερασμός των διαφόρων κρίσεων, στον οποίο να φαίνονται τόσο η εμβέλεια της κάθε γνώμης, όσο και οι αιτίες των διαφορών μεταξύ των κρίσεων. Η ομάδα των ειδικών συχνά περιλαμβάνει άτομα που προέρχονται τόσο από μέσα όσο κι απ' έξω από την επιχείρηση. Κάθε μέλος είναι ειδικός για ένα μέρος του προβλήματος, αλλά κανείς δεν είναι ειδικός για ολόκληρο το πρόβλημα. Δομημένη μέθοδος απόκτησης απαντήσεων από εμπειρογνώμονες. Αξιοποιεί την ευρεία βάση γνώσης των εμπειρογνομόνων, ενώ παράλληλα εξαφανίζει την υποκειμενική πόλωση και τον επηρεασμό των συμμετεχόντων

μέσω της ανωνυμίας .Η διαδικασία έχει επαναληπτικό χαρακτήρα και στατιστική σύνοψη στο τέλος κάθε γύρου (συνήθως 3 γύροι) , με αποτέλεσμα να προκύπτει συναίνεση (consensus) ή αποκλίνουσες απόψεις (divergent viewpoints) στο τέλος της προσπάθειας .Ο Hostrop (1975) ανακεφαλαιώνει τη σειρά των διαδικασιών (γεγονότων) που ακολουθούνται σύμφωνα με τη "DELPHI", όπως την είδε ο ιδρυτής της Helmer:

Ζητείται από τους συμμετέχοντες, που συνήθως παραμένουν άγνωστοι μεταξύ τους, να εκφέρουν τη γνώμη τους για κάποιο συγκεκριμένο ζήτημα. Η γνώμη τους δίνεται γραπτά, ώστε στη συνέχεια να προετοιμαστεί ένα ερωτηματολόγιο. Έτσι π.χ. αναφέρονται στις δραστηριότητες που προτείνονται να γίνουν για κάποιο θέμα ή κάνουν προβλέψεις για το μέλλον σχετικά με το θέμα. Στη συνέχεια ζητείται από τους μετέχοντες στην έρευνα να αξιολογήσουν το σύνολο των ιδεών (γνωμών) που εκφράστηκαν κατά το προηγούμενο στάδιο. Η αξιολόγηση αυτή γίνεται με κάποιο κριτήριο, π.χ. τη σημασία τους, την πιθανότητα επιτυχίας

Συγκεντρώνονται κατόπιν οι προτάσεις τους και ο ερευνητής τις επεξεργάζεται δίνοντάς τους ομοιόμορφη διατύπωση και διευκρινίζοντας ό,τι χρειάζεται. Στη συνέχεια κάθε μέλος της ομάδας παίρνει το νέο ερωτηματολόγιο, μαζί με την περίληψη των απαντήσεων (απόψεων) που δόθηκαν για κάθε συγκεκριμένη ερώτηση. Αν οι απόψεις ορισμένων αποτελούσαν τη μειοψηφία στο πρώτο στάδιο, τους ζητείται να επανεξετάσουν το θέμα και να αναθεωρήσουν, εφόσον γίνεται, την άποψή τους ή να αιτιολογήσουν τη διαφορετική γνώμη τους. Συγκεντρώνονται και πάλι από τον ερευνητή οι προτάσεις των μελών της ομάδας και ακολουθεί νέα επεξεργασία και νέες περιλήψεις των απαντήσεων. Κάθε μέλος της ομάδας παίρνει από τον ερευνητή νέο αντίγραφο του ερωτηματολογίου και ενημερωμένη περίληψη όλων των απαντήσεων του γύρου, ακόμη και εκείνων που έχουν διαφορετικές απόψεις. Κάθε μέλος έχει το δικαίωμα, για μια ακόμα φορά, να αναθεωρήσει

ή όχι τη γνώμη του. Τελικά, ο ερευνητής συγκεντρώνει τα αντίγραφα του ερωτηματολογίου του τελευταίου γύρου και τα ανακεφαλαιώνει για την τελική αναφορά (report). Ακολουθεί επικοινωνία του ερευνητή με κάθε μέλος της ομάδας σε μια διαδικασία ερωτήσεων, διευκρινήσεων και ανατροφοδότησης (feedback), με σκοπό την ελάττωση στο ελάχιστο της πιθανότητας λανθασμένης ερμηνείας διαφόρων απόψεων .

Αργότερα, διάφοροι ερευνητές διατύπωσαν και εφάρμοσαν διάφορες τροποποιήσεις της, ώστε να μπορεί να χρησιμοποιείται σε ποικιλία περιπτώσεων. (Linstone and Turoff, 1975). Έτσι, διαπιστώνεται σήμερα ότι υπάρχει ένας μεγάλος αριθμός ερευνών, στις οποίες χρησιμοποιήθηκε η "DELPHI" με κάποιες τροποποιήσεις. Ο Hostrop (1975) αναφέρει ότι η "DELPHI" μπορεί να δώσει στον ερευνητή περισσότερο αντικειμενικά μέσα για: (α) την εκτίμηση των διαφόρων γνώμων σχετικά με τους στόχους που τέθηκαν, (β) να δοθούν προτεραιότητες για την ιεράρχηση των αντικειμενικών σκοπών της έρευνας, (γ) να καθοριστεί ο βαθμός ομοφωνίας που υπάρχει σχετικά με τους στόχους. Γενικά, η "DELPHI" θα μπορούσε να βοηθήσει έναν ερευνητή στην εκτίμηση του "τι είναι" και "τι θα'πρεπε να είναι" κάποια κατάσταση, λαμβανομένων υπόψη των οργανωτικών συνθηκών, στόχων κ.λπ. Ο Hostrop (1975) αναφέρει ότι η μέθοδος "DELPHI" προσφέρεται σε καθηγητές και φοιτητές, για να μελετήσουν από κοινού, σε κλίμα συνεργασίας, προβλήματα της εκπαίδευσης. Στην εκπαίδευση η "DELPHI" έχει χρησιμοποιηθεί πολλές φορές, για να καθοριστούν οι στόχοι που θα'πρεπε να τεθούν για το μέλλον (Cyphert and Gant, 1971, Uhl, 1972, Hudson, 1983).

Η εφαρμογή της DELPHI

Ο βαθμός ομοφωνίας μεταξύ των ατόμων μπορεί να εξακριβωθεί (Bunning, 1979), χωρίς να υπάρξει πρόβλημα (Linstone and Turoff 1975). Ο Sweigert (1975) αναφέρεται στη σειρά κρίσεων που πάρθηκαν με τη χρήση ερωτηματολογίων ή άλλων τύπων επικοινωνίας με τη μέθοδο των επαναλήψεων κατά τη "DELPHI". Το σκεπτικό πίσω από την χρησιμοποίηση της εκτίμηση των εμπειρογνομόνων και ειδικών για την εκτίμηση και πρόβλεψη φαινομένων είναι ότι μέσα ακριβώς από την τριβή της συστηματικής μελέτης ή ενασχόλησης με ένα πεδίο, αυτοί αποκτούν θεώρηση και προβλεπτική ικανότητα επί των φαινομένων που εμπίπτουν στο πεδίο αυτό, η οποία δεν αναλύεται πάντα σε επιμέρους τεκμήρια λήψης συγκεκριμένων αποφάσεων, εκτιμήσεων και προβλέψεων (Stitt-Gohdes et al., 2004).

Αυτή η επανάληψη, κατά τον Sweigert, αποσπά την γνώμη κάθε μέλους, έτσι ώστε όλα τα μέλη της ομάδας να κάνουν μια σειρά από διαδοχικές κρίσεις. Κάθε διαδοχική κρίση γίνεται με τη βοήθεια της περίληψης των κρίσεων όλων των μελών κατά τον προηγούμενο γύρο. Αυτή η διεργασία σχεδιάζεται με σκοπό την ακριβέστερη κρίση και την αύξηση της συμφωνίας μεταξύ των μετεχόντων από γύρο σε γύρο. Η διεργασία της ανατροφοδότησης αναφέρεται στην επανάληψη. Ο κύριος λόγος, για τον οποίο σχεδιάζεται αυτή η διαδικασία επανάληψης, είναι να επιτευχθεί συμφωνία απόψεων μεταξύ των μελών της ομάδας που δε θα ήταν δυνατή σε ένα γύρο με την παραδοσιακή μέθοδο ανταλλαγής απόψεων (Sweigert, 1975). Κατά τον Dybas (1980), ο λόγος για τον οποίο σχεδιάζεται κάθε γύρος είναι να οδηγηθούν οι μετέχοντες σε μια ομοφωνία επί του θέματος που ερευνάται, η δε ανατροφοδότηση με γραπτά και με

στατιστικά δεδομένα βοηθά να κατορθωθεί αυτή η ομοφωνία. Η στατιστική μπορεί να παρουσιαστεί με πίνακες, ιστογράμματα κ.λπ. Ο MacMillan (1975) και οι Murray Turoff &

Harold A.Linstone (2002) σημειώνουν τρία σημαντικά χαρακτηριστικά της "DELPHI" που αναφέρονται στη διαδικασία της επανάληψης:

α) Την ανωνυμία των μελών που μετέχουν. Η ανωνυμία βοηθά στην ίση συμμετοχή και ευθύνη των μελών και εμποδίζει την επισήμανση της πηγής μιας άποψης που αποκλίνει. Η πίεση της ομάδας συνεπώς δεν εμποδίζει τις θέσεις των μελών. Η απόφαση (η λύση) επομένως που θα παρθεί δε θα είναι αποτέλεσμα αντίθετων καταστάσεων και αντιπαραθέσεων.

β) Την ελεγχόμενη ανατροφοδότηση που οδηγεί στον εντοπισμό των διαφορετικών απόψεων, οι οποίες είναι δυνατό να επηρεάσουν ορισμένα μέλη ή να επηρεαστούν (λόγω εξάρτησης ή για άλλους λόγους μη αντικειμενικούς) από άλλα μέλη, κατά την κατά πρόσωπο επικοινωνία, και τούτο, γιατί η επικοινωνία γίνεται με έναν ομαλό τύπο ανωνυμίας, όπου αναφέρεται η γνώμη της ομάδας και όχι η ατομική (Helmer,1966).

γ) Τη στατιστική παρουσίαση των απαντήσεων των μελών της ομάδας. Ο αριθμός των γύρων που είναι απαραίτητοι, για να υπάρξει ομοφωνία απόψεων, μπορεί να ποικίλλει, και εξαρτάται από τη φύση της μελέτης.

Ο Martino (1972) αναφέρει ότι σε πολλές περιπτώσεις δεν υπάρχει πλεονέκτημα για περισσότερους από δύο γύρους. Ο Dalkey (1968) αναφέρει ότι η πείρα έχει δείξει ότι η "DELPHI" σπάνια απαιτεί περισσότερους από τέσσερις γύρους αλλά ούτε και λιγότερους από δύο. Συγκεκριμένα κριτήρια για τον καθορισμό του αριθμού των γύρων δεν υπάρχουν, αλλά θα πρέπει να καθοριστούν από τον ερευνητή. Ο καθορισμός της ομόφωνης γνώμης των μελών της ομάδας από γύρο σε γύρο διατυπώνεται με κριτήριο την ελάττωση του εύρους των γνώμων για κάθε ερώτηση μεταξύ των γύρων. Ο Helmer (1966) σημειώνει ότι η ελάττωση των διαφορών από τον πρώτο γύρο στο δεύτερο ήταν περίπου στο 1/3 του εύρους. Ο Dybas (1980) χρησιμοποίησε ιστογράμματα, για να παραστήσει την κατανομή

συχνοτήτων των απαντώντων μεταξύ των γύρων. Η μέθοδος, με την οποία θα παρουσιαστούν τα δεδομένα που πάρθηκαν από τους διαδοχικούς γύρους της μελέτης, μπορεί να διαφέρει, εξαρτάται ωστόσο κυρίως από τη φύση των ερωτήσεων που δόθηκαν. Πάντως, κατά τον πρώτο γύρο της "DELPHI", ζητείται από τους μετέχοντες να πουν τη γνώμη τους σχετικά με το συγκεκριμένο θέμα. Από τις πληροφορίες που συλλέγονται κατά τον πρώτο γύρο, ο ερευνητής μπορεί να συνθέσει, αλλά και να αφαιρέσει, το άσχετο υλικό, για να ετοιμάσει τέλος ένα ερωτηματολόγιο για τον δεύτερο γύρο (Hostrop, 1975). Τα δεδομένα δηλ. από τον πρώτο γύρο συλλέγονται και επανυποβάλλονται στην ομάδα με στατιστικό ή με λεκτικό τρόπο ως δεύτερος γύρος. Ο δεύτερος γύρος δίνει στα μέλη τη δυνατότητα, αφού λάβουν υπόψη τις γνώμες της ομάδας, να επαναξιολογήσουν τις αντιδράσεις που είχαν κατά τον πρώτο γύρο. Κάθε μέλος απαντά στις ίδιες ερωτήσεις που είχε και ο πρώτος γύρος. Μπορεί να διατηρήσει την ίδια άποψη (θέση) που είχε ή να την αλλάξει κάτω από το φως των νέων πληροφοριών. Ο Sweigert (1975) αναφέρει ότι αν κάποια μέλη επιθυμούν να πάρουν θέση διαφορετική από εκείνη που επικρατεί στην ομάδα, πρέπει να την υποστηρίξουν. Αντίθετα, ο Dybas (1980) υποστηρίζει ότι ο καθένας πρέπει να υποστηρίξει την επιλογή του ανεξάρτητα από το αν η θέση που παίρνει συγκλίνει ή αποκλίνει. Προφορικά σχόλια που δίνονται από τον ερευνητή κατά τον επόμενο γύρο μπορούν να βοηθήσουν τον καθένα που συμμετέχει στην ανάλυση των διαφόρων απόψεων. Μετά τη λήψη των απαντήσεων από το δεύτερο γύρο, ο ερευνητής επεξεργάζεται τα δεδομένα και ανατροφοδοτεί τα μέλη της ομάδας για τον τρίτο γύρο, εάν αυτός κρίνεται απαραίτητος. Γραπτά σχόλια, που συγκεντρώθηκαν κατά τον δεύτερο γύρο, ανακεφαλαιώνονται μαζί με τη στατιστική παρουσίαση των απαντήσεων των μελών.

Ανατροφοδοτούνται σε κάθε μέλος οι πληροφορίες, ώστε να τις έχουν υπόψη τους, και μαζί με τις προσωπικές τους απόψεις να οδηγηθούν και πάλι στην υποστήριξη των νέων απόψεών τους κατά τον τρίτο γύρο. Ο ερευνητής πρέπει να καθορίσει αν ο βαθμός

ομοφωνίας που επιτεύχθηκε είναι αρκετός, ενώ ταυτόχρονα πρέπει να λάβει υπόψη του και τα άτομα που υποστηρίζουν αποκλίνουσες απόψεις. Τέλος, τα δεδομένα από τον τρίτο γύρο πινακοποιούνται, και αποφασίζεται κατά πόσο χρειάζεται και άλλος γύρος. Αν ναι, ακολουθείται η ίδια διαδικασία, όπως και στον τρίτο γύρο. Αν όχι, συντάσσεται η τελική αναφορά και μοιράζεται σε κάθε μέλος. Επειδή δεν υπάρχει προσωπική επαφή μεταξύ των μελών, ο ερευνητής έχει την ευθύνη να γεφυρώσει το κενό (Dalkey, 1968). Μπορεί να επικοινωνήσει (κατά τη διαδικασία) με τα μέλη της ομάδας στις περιπτώσεις που είναι απαραίτητο να διασαφηνίσει κάποια σκοτεινά σημεία. Τα μέλη της ομάδας, εξάλλου, είναι επίσης ελεύθερα να επικοινωνήσουν με τον ερευνητή οποιαδήποτε στιγμή, για να ξεκαθαρίσουν διαδικαστικές απορίες που μπορεί να παρουσιαστούν. Στην ουσία, η επανάληψη των απαντήσεων παρέχει μια διεργασία για εξευγένιση της λύσης. Η τεχνική "DELPHI", επομένως, διευκολύνει την ομοφωνία γνώμών χωρίς την προσωπική αντιπαράθεση των μελών. Η ακρίβεια των απαντήσεων τείνει να αυξηθεί με την επανάληψη (Dalkey, 1968). Η μεγαλύτερη τάση για ομοφωνία αίνεται να συμβαίνει κυρίως μεταξύ πρώτου και δευτέρου γύρου. Η απόφαση του ερευνητή να σταματήσει τη διαδικασία βασίζεται στην ανάλυση των πληροφοριών που του δίδονται υπό το φως των αντικειμενικών σκοπών της μελέτης .

Όπως συμβαίνει και σε άλλες ερευνητικές μεθόδους, η Delphi έχει δεχθεί κατά καιρούς κριτική αναφορικά με την εγκυρότητα και την αξιοπιστία των αποτελεσμάτων της. Η εγκυρότητα και η αξιοπιστία της μεθόδου Delphi διαφέρουν ως προς τον τρόπο που ορίζονται σε σχέση με άλλες ποσοτικές έρευνες, καθώς τα ερωτηματολόγια είναι συχνά γενικής φύσης και ανοικτού τύπου. Σχετικά με την αξιοπιστία, οι Watson et al αναφέρουν ότι δεν υπάρχουν έως τώρα ενδείξεις ότι, εάν οι ίδιες πληροφορίες δοθούν σε μια άλλη ομάδα εμπειρογνομόνων, θα προκύψουν τα ίδια αποτελέσματα. Η διαδικασία του ελέγχου επαναληψιμότητας (test-retest) θα μπορούσε να εφαρμοστεί θεωρητικά. Η ομάδα όμως των

εμπειρογνομόνων δεν είναι τόσο ανεκτική στη διαδικασία της επανάληψης. Η αξιοπιστία μπορεί επίσης να υπολογιστεί ανάμεσα στα ερωτηματολόγια, π.χ. των δύο πρώτων γύρων. Εάν προκύπτει ένα ικανό ποσοστό συναίνεσης στο δεύτερο γύρο, ενδεχομένως ο ερευνητής έχει ομαδοποιήσει σωστά το νόημα που κρύβεται πίσω από τις απαντήσεις του πρώτου γύρου. Για το θέμα της εγκυρότητας, η Goodman υποστηρίζει ότι μόνο εάν τα μέλη της ομάδας προέρχονται από αντι-προσωπευτικές ομάδες του πληθυσμού και ασχολούνται με συγκεκριμένα γνωστικά αντικείμενα, η εγκυρότητα μπορεί να θεωρηθεί ότι υφίσταται. Εξ άλλου και η ίδια η έννοια της εγκυρότητας έχει αμφισβητηθεί, καθώς ένα συμπέρασμα μπορεί να περιγράψει μια κατάσταση του παρόντος με ορθό τρόπο, αλλά σε μια μελλοντική διαφοροποίηση της πραγματικότητας το συμπέρασμα αυτό έχει αξία μόνον ως «ιστορία». Συνεπώς, κάθε γενίκευση αποτελεί μια υπόθεση υπό διαρκή έλεγχο και όχι ένα διαχρονικό συμπέρασμα.

ΠΛΕΟΝΕΚΤΗΜΑΤΑ ΚΑΙ ΠΕΡΙΟΡΙΣΜΟΙ

Η μέθοδος Delphi χρησιμοποιείται ευρέως στην έρευνα, όταν τα δεδομένα είναι ανεπαρκή ή ακατάλληλα για να αξιοποιηθούν μέσω άλλων ερευνητικών μεθόδων. Επειδή η μέθοδος Delphi είναι σχετικά απλή στη χρήση της, δεν απαιτούνται προχωρημένες γνώσεις στατιστικής και μαθηματικών κατά το σχεδιασμό, την εφαρμογή και την ανάλυσή της. Ένα επί πλέον πλεονέκτημα της συγκεκριμένης μεθόδου είναι η δυνατότητα που παρέχεται στους συμμετέχοντες να εκφράζουν την προσωπική τους άποψη μέσα από δεδομένα που αναδύονται από την εμπειρία όλων των μελών της ομάδας. Στη συμπλήρωση των ερωτηματολογίων δεν υπάρχουν γεωγραφικοί περιορισμοί, αφού συνήθως συμπληρώνονται μέσω παραδοσιακού ή ηλεκτρονικού ταχυδρομείου. Επίσης, οι εμπειρογνώμονες δεν περιορίζονται στο χρονικό όριο που θα διαρκούσε μια συνάντηση με τον ερευνητή, αλλά συμμετέχουν στα διάφορα στάδια της έρευνας με ασύγχρονο τρόπο. Έτσι, είναι σε θέση να

εξετάσουν διεξοδικότερα το υπό μελέτη θέμα και να παρέχουν τη συμβολή τους χωρίς χρονικούς περιορισμούς ή επηρεασμό από τα μέλη της ομάδας που έχουν ισχυρή προσωπικότητα. Όπως αναφέρει ο McKenna, πολλοί επιστήμονες υποστηρίζουν ότι η μέθοδος Delphi παρέχει οικονομία χρόνου και χρήματος, αν και η αντίθετη άποψη έχει επίσης διατυπωθεί . Όπως συμβαίνει σε όλα τα ερευνητικά εγχειρήματα, η διάρκεια ολοκλήρωσης και το κόστος εξαρτώνται από την κλίμακα μεγέθους του δείγματος. Στην έρευνα Delphi, εξαρτάται επί πλέον από τον αριθμό των γύρων που θα διεξαχθούν έως την επίτευξη της μέγιστης δυνατής συναίνεσης.

Επειδή η μέθοδος Delphi απαιτεί πολλαπλούς γύρους ερωτηματολογίων, αυξάνεται το ποσοστό των εγκαταλείψεων έως τη συμπλήρωση του τελευταίου ερωτηματολογίου. Αυτός είναι ένας σημαντικός περιορισμός για την αξιοπιστία της έρευνας, όπως για παράδειγμα στην έρευνα της Moreno- Casbas, στην οποία το ποσοστό ανταπόκρισης ήταν της τάξης του 21%. Επίσης, η Delphi έχει κατηγορηθεί ότι ωθεί προς τη συναίνεση και δεν επιτρέπει στους συμμετέχοντες να συζητούν μεταξύ τους για τα θέματα που εγείρονται, ενώ μειώνει τη θετική επενέργεια της αλληλόδρασης μιας ομαδικής συνάντησης, κατά την οποία η ανταλλαγή πληροφοριών βοηθά στην αναγνώριση των λόγων της διαφωνίας. Επί πλέον, έχει αναφερθεί ότι η παρουσίαση της άποψης των ειδικών δεν αποτελεί αδιαμφισβήτητη αλήθεια. Στην απουσία όμως της αδιαφιλονίκητης γνώσης, η άποψη των ειδικών, όπως παρέχεται μέσα από τη μέθοδο Delphi, είναι ισχυρότερη από άλλες μεθόδους συναίνεσης, για παράδειγμα τις ομάδες εστιασμένης συζήτησης.

Δομημένη αναλογία

Μια δομημένη προσέγγιση που περιλαμβάνει μια ομάδα εμπειρογνομόνων όπως προτείνεται από τους Green and Armstrong . Η ιδέα είναι παρόμοια με εκείνη της DELPHI , ωστόσο το έργο πρόβλεψης ολοκληρώνεται μέσω την εξέταση αναλογιών . Πρώτα απ 'όλα διορίζεται ένας συντονιστής (facilitator). Στη συνέχεια, η δομημένη προσέγγιση περιλαμβάνει τα ακόλουθα βήματα .Αρχικά συγκροτείται μια ομάδα εμπειρογνομόνων, οι οποίοι είναι πιθανό να έχουν εμπειρία σε ανάλογες καταστάσεις .Οι εργασίες / προκλήσεις διανέμονται στους εμπειρογνώμονες

Οι ειδικοί καλούνται να αναγνωρίσουν και να περιγράψουν όσες αναλογίες μπορούν , κατηγοριοποιούν τις ομοιότητες και τις διαφορές της κάθε αναλογίας με την κατάσταση , έπειτα βαθμολογείται η ομοιότητα κάθε αναλογίας με την εγγύτητα και συγγένεια σε μια κλίμακα . Οι προβλέψεις προέρχονται από τον συντονιστή με τη χρήση ενός κανόνα σκετ . Αυτό μπορεί να είναι ένας σταθμισμένος μέσος όρος , όπου τα βάρη μπορεί να αποδίδονται από τα αποτελέσματα κατάταξης για κάθε μία αναλογία από τους εμπειρογνώμονες . Ομοίως, με την προσέγγιση της Delphi , η ανωνυμία των εμπειρογνομόνων θα μπορούσε να είναι ένα πλεονέκτημα που δεν καταστέλλει τη δημιουργικότητα , αλλά θα μπορούσε να εμποδίσει τη συνεργασία . Οι Green και Armstrong δεν βρήκαν κανένα κέρδος στην συνεργασία των εμπειρογνομόνων για τα αποτελέσματά τους . Ένα σημαντικό εύρημα ήταν ότι οι εμπειρογνώμονες με πολλαπλές αναλογίες (περισσότερες από δύο) και οι οποίοι είχαν εμπειρία με τις αναλογίες , αποδίδουν πιο ακριβείς προβλέψεις . Η μέθοδος των αναλογιών είναι μια διαδικασία δημιουργίας κριτικών προβλέψεων που περιλαμβάνει τρία στάδια (Lee et al., 2007)

- A. Την ανάκληση των αναλογιών
- B. Τις κριτικές προβλέψεις που προκύπτουν από τις αναλογίες με βάση την ομοιότητα τους με το τρέχον πρόβλημα
- C. Τις προσαρμογές στις κριτικές προβλέψεις που αφορούν τις ειδικές συνθήκες και καταστάσεις του προβλήματος

Οι αναλογίες είναι χρήσιμες σε πολύπλοκες περιπτώσεις και μπορούν να βελτιώσουν τις προβλέψεις που γίνονται με απλή κρίση. Η μέθοδος των δομημένων αναλογιών (Green and Armstrong 2005) εξελίσσεται σε τέσσερα βήματα :

- A. Περιγραφή της τρέχουσας κατάστασης
- B. Περιγραφή και προσδιορισμό των αναλογιών
- C. Αποτίμηση της ομοιότητας με παρελθούσες καταστάσεις
- D. Παραγωγή των προβλέψεων

Αιτιοκρατικό Μοντέλο

Το αιτιοκρατικό μοντέλο υποθέτει ότι η υπό πρόβλεψη μεταβλητή έχει μία αιτιοκρατική σχέση με μία άλλη, ανεξάρτητη μεταβλητή. Το σύστημα είναι μια συνάρτηση της μορφής $y=f(x_1, x_2, \dots, x_n)$. Η σχέση αυτή προφανώς είναι μοναδική. Παράδειγμα τέτοιου μοντέλου θα μπορούσαμε να θεωρήσουμε ως εξαρτημένη μεταβλητή το ύψος των πωλήσεων ενός προϊόντος και ως ανεξάρτητη μεταβλητή την τιμή πώλησης .

Σε σύγκριση με τα μοντέλα χρονοσειρών , στα αιτιοκρατικά μοντέλα η συνάρτηση που περιγράφει το σύστημα ακολουθεί μια συγκεκριμένη διαδικασία. Αρχικά , παρατηρείται η τιμή του y και συσχετίζεται με τις αντίστοιχες τιμές των παραμέτρων X που θεωρούνται δεδομένες . Στην συνέχεια προσδιορίζεται το είδος και η ακριβή μορφή της σχέσης (συνάρτηση f) που συνδέει την έξοδο με τις εισόδους του συστήματος .

Τα αιτιοκρατικά μοντέλα μπορούν να χρησιμοποιηθούν σε πολλές εφαρμογές. Χαρακτηριστικά παραδείγματα είναι η εθνική οικονομία, η αγορά μίας επιχείρησης και άλλα. Η πρόβλεψη εκφράζεται ως συνάρτηση κάποιου συγκεκριμένου αριθμού παραγόντων που επηρεάζουν την τελική τιμή της. Δεν είναι απαραίτητο να υπάρχει χρονική εξάρτηση. Αναπτύσσεται ένα μοντέλο το οποίο καθιστά ευκολότερη την κατανόηση των συνθηκών και δίνει τη δυνατότητα πρόβλεψης μελλοντικής τιμής κάποιου μεγέθους μέσω διαφόρων συνδυασμών τιμών των ανεξάρτητων μεταβλητών.

Μειονέκτημα αυτών των μεθόδων είναι το μεγάλο πλήθος δεδομένων καθώς υπάρχει απαίτηση δεδομένων σχετικών με τις ανεξάρτητες μεταβλητές , επίσης η πρόβλεψη με βάση αιτιοκρατικά

μοντέλα προϋποθέτει πρόβλεψη και των ανεξάρτητων μεταβλητών, κάτι που συνεπάγεται και αυξημένο κόστος εφαρμογής.

Στα αιτιοκρατικά μοντέλα ανήκουν οι μέθοδοι παλινδρόμησης και οι οικονομετρικές μέθοδοι .

Μέθοδοι Πρόβλεψης Συνεχούς Ζήτησης

Μοντέλο Χρονοσειρών

Η ανάλυση χρονοσειρών (time series analysis) ασχολείται με τη διερεύνηση της διαχρονικής συμπεριφοράς των τιμών μιας μεταβλητής, οι παρατηρήσεις της οποίας προέρχονται από χρονοσειρά. Η πρόβλεψη των μελλοντικών τιμών της μεταβλητής σύμφωνα με την ανάλυση χρονοσειρών μπορεί να προέλθει από τις παρακάτω κατηγορίες μεθόδων προβλέψεων:

- Μέθοδοι Εξομάλυνσης
- Διάσπαση χρονοσειρών
- Ανάλυση ARIMA

Το μοντέλο των χρονοσειρών σε αντίθεση με το αιτιοκρατικό μοντέλο δεν αναλύει τη σχέση που υπάρχει μεταξύ της υπό εξέταση μεταβλητής και άλλων ανεξάρτητων μεταβλητών αλλά θεωρεί το συνολικό σύστημα ανεπηρέαστο και δεν ασχολείται καθόλου με τους παράγοντες που το επηρεάζουν. Γι αυτόν τον λόγο και η πρόβλεψη για το μέλλον βασίζεται αποκλειστικά σε ιστορικά δεδομένα της υπό εξέτασης μεταβλητής και σε σφάλματα που προκύπτουν από αυτές. Ο σκοπός της χρήσης τέτοιων μεθόδων πρόβλεψης είναι η εύρεση ενός μοντέλου που ακολουθεί η αντίστοιχη παρατήρηση και η προέκτασή του στο μέλλον.

Και τα δύο μοντέλα έχουν πλεονεκτήματα και μειονεκτήματα σε διάφορες εφαρμογές. Η επιλογή τους κρίνεται εν γένει από τα διαθέσιμα δεδομένα, τις συνθήκες που επικρατούν στην αντίστοιχη περίπτωση και κυρίως από τον στόχο προς επίτευξη μέσω της πρόβλεψης.

Στην κοινότητα των προβλέψεων, δίνεται ιδιαίτερη βαρύτητα σε δύο κύριους τύπους στατιστικών παρατηρήσεων: στα διαστρωματικά στοιχεία (cross – sectional data) και στις χρονολογικές σειρές ή χρονοσειρές δεδομένων (time series).

Τα cross – sectional data αφορούν μετρήσεις μεταβλητών για διάφορες οικονομικές μονάδες όπως άτομα, χώρες, επιχειρήσεις κ.λπ. Δεδομένα αποτελούνται από πολλές παρατηρήσεις για ένα συγκεκριμένο μέγεθος για το ίδιο χρονικό διάστημα ενώ οι χρονοσειρές αποτελούνται από μία αλληλουχία διαχρονικών παρατηρήσεων του ίδιο μεγέθους.

Λόγω της φύσης του αντικείμενου μελέτης της παρούσης διπλωματικής εργασίας, θα δοθεί έμφαση στην ανάλυση των χρονοσειρών καθώς αποτελούν δομικό στοιχείο της παρούσης μελέτης αλλά και γενικότερα της μεθοδολογίας των προβλέψεων.

Οι μέθοδοι χρονοσειρών κατηγοριοποιούνται με της σειρά τους ως εξής:

Αντικείμενο των μεθόδων αποσύνθεσης είναι ο διαχωρισμός κύριων χαρακτηριστικών των χρονοσειρών και η απομόνωσή τους. Τα κύρια αυτά χαρακτηριστικά των χρονοσειρών είναι: η τάση, ο κύκλος, η εποχιακότητα και η τυχαιότητα.

Εκτίμηση τάσης : Υπολογίζονται αρχικά οι κινητοί μέσοι όροι αντίστοιχου μήκους με την εποχιακότητα που παρουσιάζει η χρονοσειρά και συνεχίζεται με όμοιο τρόπο η διαδικασία με

τους κινητούς μέσους όρους μήκους 2. Με αυτόν τον τρόπο υπολογίζεται η εκτίμηση της τάσης η οποία είναι οι κεντρικοί κινητοί μέσοι όροι.

Υπολογισμός Λόγων Εποχιακότητας (Λ.Ε.) : Υπολογίζονται ξεχωριστά ανά περίοδο (μήνα, τρίμηνο κτλ) ανάλογα με τη φύση των ιστορικών στοιχείων ως ο λόγος της πραγματικής τιμής της χρονοσειράς προς την αντίστοιχη τιμή που έχει προκύψει από τους κεντρικούς κινητούς μέσους όρους, ακολουθεί κανονικοποίηση τους σύμφωνα με το άθροισμα που πρέπει να έχουν και προκύπτει η αποεποχικοποιημένη χρονοσειρά από την αρχική χρονοσειρά διαιρώντας με τους αντίστοιχους εποχιακούς δείκτες την κάθε παρατήρηση.

Μπορούμε να θεωρήσουμε δύο κύριες κατηγορίες μεθόδων πρόβλεψης. Στην **πρώτη κατηγορία** έχουμε τις *ποσοτικές* οι οποίες χρησιμοποιούνται, όταν έχουμε διαθέσιμη πληροφορία για το παρελθόν (επαρκή δεδομένα), όταν μπορεί να γίνει ποσοτικοποίηση της πληροφορίας (αριθμητικά δεδομένα) και τέλος όταν θεωρούμε ότι το πρότυπο συμπεριφοράς θα διατηρηθεί και στο μέλλον.

Στη **δεύτερη κατηγορία** έχουμε τις *κριτικές (judgmental)* ή *ποιοτικές (qualitative)* μεθόδους οι οποίες χρησιμοποιούνται, όταν έχουμε διαθέσιμες λίγες ή καθόλου πληροφορίες για το παρελθόν. Σε αυτές τις περιπτώσεις απαιτείται εμπειρία, γνώση και κριτική ικανότητα. Χρησιμοποιούνται, κυρίως, σε συνδυασμό με τις ποσοτικές μεθόδους, ενώ δεν απαντούν μόνες σε αριθμητικές προβλέψεις. Οι ποσοτικές μέθοδοι προβλέψεων έχουν ως πρωταρχικό τους σκοπό τη διερεύνηση του τρόπου δημιουργίας των τιμών της υπό εξέταση μεταβλητής. Αν ο σκοπός αυτός επιτευχθεί με επιτυχία, τότε η χρησιμοποιούμενη μέθοδος μπορεί να εφαρμοστεί για τη διαμόρφωση προβλέψεων που αφορούν στη μελλοντική εξέλιξη των τιμών της συγκεκριμένης μεταβλητής. Αυτό σημαίνει ότι οι προβλέψεις (forecasts) δημιουργούνται από

την εφαρμογή εκείνης της μεθόδου που ερμηνεύει με τον καλύτερο τρόπο τη συμπεριφορά των τιμών της μεταβλητής [23].

Η διαμόρφωση των προβλέψεων στηρίζεται κατά κανόνα σε δύο βασικές προϋποθέσεις. Η πρώτη αφορά στη χρησιμοποιούμενη ποσοτική μέθοδο και η δεύτερη στον τρόπο συμπεριφοράς της μεταβλητής. Με άλλα λόγια, η επιλεγόμενη μέθοδος θα πρέπει να προσδιορίζει με τον καλύτερο τρόπο, με τη στατιστική έννοια του όρου, τη συμπεριφορά των τιμών της μεταβλητής. Όσο καλύτερα αναγνωρίζεται ο τρόπος δημιουργίας των τιμών της μεταβλητής, τόσο καλύτερες προβλέψεις αναμένεται να διαμορφώνονται. Για να ισχύσει, όμως, κάτι τέτοιο, θα πρέπει, επιπροσθέτως, η υπάρχουσα δομή του τρόπου συμπεριφοράς των τιμών της μεταβλητής να παραμένει σταθερή ή περίπου σταθερή και στο άμεσο μέλλον.

Με βάση τη στατιστική ανάλυση, η συμπεριφορά των τιμών μιας μεταβλητής προσδιορίζεται από δύο μέρη, που είναι το **συστηματικό** (systematic) και το **τυχαίο** (random). Το συστηματικό μέρος αναφέρεται σε όλους εκείνους τους παράγοντες που μπορούν να καθοριστούν και οι οποίοι ερμηνεύουν τον τρόπο συμπεριφοράς της μεταβλητής. Αντίθετα, το τυχαίο μέρος αφορά όλους εκείνους τους τυχαίους και απρόσμενους παράγοντες οι οποίοι, αν και επηρεάζουν τις τιμές της μεταβλητής, δεν μπορούν, ωστόσο, να προσδιοριστούν. Αυτό σημαίνει πως, όταν αναφερόμαστε στον προσδιορισμό του τρόπου συμπεριφοράς των τιμών μιας μεταβλητής, εννοούμε αποκλειστικά τον καθορισμό του συστηματικού μέρους. Έτσι, οι προβλέψεις που προέρχονται από τη χρησιμοποίηση των ποσοτικών μεθόδων, διαμορφώνονται με βάση το συστηματικό μέρος και όχι το τυχαίο, η συμπεριφορά [24] του οποίου δε μπορεί να προσδιοριστεί, αφού δεν παρατηρείται.

Ποιοτικά χαρακτηριστικά χρονοσειρών - Στατιστικοί δείκτες & δείκτες σφαλμάτων

Οι χρονοσειρές είναι ιστορικά δεδομένα που απαρτίζονται από διαδοχικές παρατηρήσεις μέσα σε ένα χρονικό διάστημα. Οι παρατηρήσεις γίνονται ανά σταθερό χρονικό βήμα και μπορούν να είναι ετήσιες, τριμηνιαίες, μηνιαίες, εβδομαδιαίες, ημερήσιες κτλ. Τα ποιοτικά χαρακτηριστικά αυτών των παρατηρήσεων είναι τα εξής [23]:

- Στασιμότητα (Stationary), όταν οι τιμές κυμαίνονται γύρω από μία μέση τιμή.
- *Τάση* (Trend , T) ορίζεται ως μια μακροπρόθεσμη μεταβολή του μέσου επιπέδου των τιμών της χρονοσειράς ,Μπορεί να είναι αύξουσα, φθίνουσα είτε μηδενική τάση, ανάλογα με την πορεία των στοιχείων της χρονοσειράς σε μακροπρόθεσμο ορίζοντα. Η τάση κατηγοριοποιείται ακόμη ως γραμμική και μη γραμμική (παραδείγματα μη γραμμικής τάσης: εκθετική, λογαριθμική). Εφόσον αναφερόμαστε σε μακροπρόθεσμη μεταβολή, είναι σημαντικό να διευκρινίζεται ο χρονικός ορίζοντας της μεταβολής που θεωρείται ως μακροπρόθεσμος, καθώς είναι δυνατόν το ίδιο τμήμα της χρονοσειράς να θεωρείται ότι παρουσιάζει τάση για σχετικά μικρή περίοδο, ενώ για μεγαλύτερη χρονική περίοδο θα παρατηρούσαμε κυκλική διακύμανση των δεδομένων.όταν υπάρχει μια μακροπρόθεσμη αύξηση ή μείωση του επιπέδου των τιμών.
- Εποχικότητα (Seasonal , S) αποτελεί μια περιοδική διακύμανση με σταθερό μήκος, μικρότερο του έτους. Είναι εμφανής η διαφορά της εποχικότητας από την κυκλικότητα στο μήκος της περιόδου (μικρότερη έναντι μεγαλύτερης του έτους), όπως και στη σταθερότητα της περιόδου (σταθερή έναντι όχι απαραίτητα σταθερής)

- Κυκλικότητα (Cyclical , C)ορίζεται ως μια κυματοειδής μεταβολή οφειλόμενη σε ειδικές εξωγενείς συνθήκες, η οποία παρουσιάζεται σε περιόδους όχι απαραίτητα σταθερής διάρκειας και με μήκος κατά κανόνα μεγαλύτερο του έτους. Οι εξωγενείς συνθήκες είναι κυρίως οι οικονομικές συνθήκες των διαφόρων περιόδων, στις οποίες παρουσιάζονται διαδοχικές άνοδοι και υφέσεις και ονομάζονται επιχειρηματικοί κύκλοι όταν οι τιμές αυξομειώνονται, αλλά όχι σε σταθερές περιόδους.
- Ασυνέχειες (outliers), οι οποίες οφείλονται είτε σε ασυνήθιστες τιμές, είτε σε αλλαγή επιπέδου (level shift).
- Τυχασιότητα (Irregular-Random , R), όταν έχουμε διακυμάνσεις λόγω τυχαίων γεγονότων. Αναφέρεται σε τυχαίες και μη προβλέψιμες διακυμάνσεις των δεδομένων μιας χρονοσειράς. Είναι η διακύμανση που παραμένει στη χρονοσειρά έπειτα από την απομόνωση της τάσης, της κυκλικότητας και της τυχασιότητας. Οι διακυμάνσεις τυχασιότητας ενδέχεται να είναι ασυνέχειες με παροδικό ή μόνιμο χαρακτήρα

Οι δύο απλούστερες διατυπώσεις της αποσύνθεσης είναι η προσθετική και η πολλαπλασιαστική, με τη μορφή που παρουσιάζεται στις επόμενες δύο σχέσεις:

$$Y_t = S_t \cdot T_t \cdot C_t \cdot R_t$$

$$Y_t = S_t + T_t + C_t + R_t$$

Η μέθοδος των κινητών μέσων όρων δεν παρέχει πληροφορίες σχετικά με την εποχικότητα και την τυχαιότητα της χρονοσειράς τις οποίες απομακρύνει από αυτή για την παραγωγή της σειράς τάσης-κύκλου. Υπάρχουν τέσσερα είδη κινητών μέσων όρων για την παραγωγή της εξομαλυμένης σειράς. Με την αύξηση του μήκους n επιτυγχάνεται και μεγαλύτερη εξομάλυνση των δεδομένων, ωστόσο πολύ μεγάλη τιμή του μήκους οδηγεί στον υπολογισμό της μέσης τιμής των δεδομένων της χρονοσειράς και δεν διακρίνονται η τάση και ο κύκλος που μας ενδιαφέρουν, ενώ επιπλέον αυξάνεται πολύ ο αριθμός των κενών τιμών στην αρχή και στο τέλος της παραγόμενης σειράς. Να σημειωθεί ότι στην περίπτωση έντονης εμφάνισης εποχικότητας προτιμάται η χρήση μήκους ίσου (ή μεγαλύτερου) με το μήκος της περιόδου της εποχικότητας.

Ο απλός κινητός μέσος όρος (ΚΜΟ) πρόκειται για τον απλό μέσο όρο n τιμών (όπου n : περιττός για τη διατήρηση συμμετρίας στους υπολογισμούς) της αρχικής χρονοσειράς γύρω από την εκάστοτε παρατήρηση. Υπολογίζοντας τον απλό κινητό μέσο όρο για όλες τις παρατηρήσεις της χρονοσειράς προκύπτει η επιθυμητή σειρά τάσης-κύκλου.

Ο σταθμισμένος κινητός μέσος όρος (ΣΚΜΟ) αποτελεί παραλλαγή του απλού ΚΜΟ με συμμετοχή των γειτονικών παρατηρήσεων στον υπολογισμό της σειράς τάσης-κύκλου με άνισα βάρη. Το άθροισμα από τα βάρη πρέπει να ισούται με τη μονάδα, δίνεται προσοχή ώστε να είναι συμμετρικά ως προς την εκάστοτε (τρέχουσα) παρατήρηση, ενώ η επιλογή τους γίνεται με στόχο να δοθεί μεγαλύτερη βαρύτητα σε παρατηρήσεις πλησίον της εκάστοτε παρατήρησης και μικρότερη στις πιο μακρινές παρατηρήσεις. Η χρήση του σταθμισμένου κινητού μέσου όρου οδηγεί σε περισσότερο εξομαλυμένη σειρά τάσης- κύκλου.

Στο διπλό κινητό μέσο όρο (ΔΚΜΟ) με διπλή εφαρμογή του απλού ΚΜΟ, με ίσα ή και άνισα μήκη, οδηγούμαστε σε διπλή εξομάλυνση, αλλά και στην ύπαρξη περισσότερων κενών τιμών - συγκριτικά πάντα με τον απλό κινητό μέσο όρο (ΚΜΟ). Ακόμη, η εφαρμογή του ΔΜΚΟ ισοδυναμεί με την εφαρμογή του ΣΚΜΟ με ορισμένα βάρη (εξαρτώμενα από την κάθε περίπτωση).

Ο κεντρικός κινητός μέσος όρος (ΚΚΜΟ) αποτελεί συνδυασμό του απλού και του διπλού κινητού μέσου όρου και χρησιμεύει για τον υπολογισμό κινητού μέσου όρου άρτιου μήκους. Η εφαρμογή του κεντρικού κινητού μέσου όρου ισοδυναμεί με ένα διπλό κινητό μέσο όρο $2 \times n$ (όπου n το μήκος του απλού ΚΜΟ), καθώς και με ένα σταθμισμένο κινητό μέσο όρο με ορισμένα βάρη.

Κλασσική Μέθοδος Αποσύνθεσης

Αποτελεί την πιο απλή διαδικασία για την απομόνωση των τεσσάρων συνιστωσών μιας χρονοσειράς. Τα πέντε βασικά βήματα που πρέπει να ακολουθήσουμε είναι τα εξής :

1ο Βήμα

Ο υπολογισμός ενός κινητού μέσου όρου βασίζεται στο μήκος της εποχικότητας.

Δεν εμπεριέχεται στον κινητό μέσο όρο η εποχικότητα ενώ περιέχει πολύ μικρή ή μηδενική τυχαιότητα, δεδομένου ότι η τυχαιότητα φανερώνει τυχαίες διακυμάνσεις γύρω από τον μέσο όρο παρατηρήσεων. Οι τιμές των κινητών μέσων όρων που υπολογίζονται με αυτόν τον τρόπο είναι εξομαλυμένες και δίνουν καλή εκτίμηση της συμπεριφοράς της χρονοσειράς.

$$KMO(n) = T \times C$$

KMO(n) είναι κινητός μέσος όρος μήκους n και T,C οι συνιστώσες τάσης και κύκλου .

2ο Βήμα

Διαίρεση των πραγματικών δεδομένων με τις αντίστοιχες τιμές των κινητών μέσων όρων . Με την διαδικασία αυτή προκύπτουν οι λόγοι εποχικότητας , οι οποίοι εμπεριέχουν τυχαιότητα .

$$\frac{Y}{KMO(n)} = \frac{T \cdot C \cdot S \cdot R}{T \cdot C} = S \cdot R$$

3ο Βήμα

Σε αυτό το σημείο , σκοπός του βήματος είναι η απαλοιφή της τυχαιότητας από τους λόγους εποχικότητας του προηγούμενου βήματος .Αρχικά υπολογίζονται οι μέσοι όροι των αντίστοιχων λόγων εποχικότητας , οι οποίοι αποτελούν τους δείκτες εποχικότητας της αρχικής χρονοσειράς . Πολλές φορές είναι απαραίτητη η κανονικοποίηση ώστε το άθροισμα τους να ισούται με το μήκος της εποχικότητας χωρίς να χρησιμοποιούνται οι ακραίες τιμές .Σε περίπτωση που η χρονοσειρά περιέχει αρκετή τυχαιότητα ή ασυνήθιστες τιμές , προτείνεται ο υπολογισμός των ενδιάμεσων μέσω όρων , ώστε να προκύψουν οι δείκτες εποχικότητας .

4ο Βήμα

Διαίρεση των πραγματικών δεδομένων με τους αντίστοιχους δείκτες εποχικότητας για την εύρεση της αποεποχικοποιημένης σειράς . Η σειρά αυτή περιέχει τις υπόλοιπες τρεις συνιστώσες , την τάση , τον κύκλο και την τυχαιότητα :

$$\frac{Y}{S} = \frac{T \cdot C \cdot S \cdot R}{S} = T \cdot C \cdot R$$

5ο Βήμα

Ενδιαφέρον έχει η απαλοιφή και της τυχειότητας από την ήδη αποεποχικοποιημένη σειρά. Υπολογίζεται, ο κινητός μέσος όρος μήκους $n = 3$ ή $n = 6$ παρατηρήσεων της αποεποχικοποιημένης σειράς, ώστε να παραχθεί μια ικανοποιητική και ακριβής σειρά τάσης-κύκλου. Για καλύτερα αποτελέσματα συνίσταται η χρήση διπλού κινητού μέσου όρου, ΔΚΜΟ(3x3). Στην περίπτωση που ζητείται η συνιστώσα της τυχειότητας, υπολογίζεται με βάση τη σχέση:

$$\frac{T \cdot C \cdot R}{KMO(3 \times 3)} = \frac{T \cdot C \cdot R}{T \cdot C} = R$$

Οι περισσότερες μέθοδοι προβλέψεων στηρίζονται, σε κάποιο βαθμό, στις βασικές έννοιες της στατιστικής ανάλυσης. Για αυτό το λόγο θα κάνουμε μια σύντομη αναφορά στους πιο βασικούς στατιστικούς δείκτες .

Μέση τιμή (mean) : $MEAN = \frac{\sum_{i=1}^n x_i}{n}$

Διάμεσος (median) : η μεσαία παρατήρηση αν (n) περιττός , ο μέσος των δύο “μεσαίων” παρατηρήσεων αν (n) άρτιος

Mean Absolute deviation : $MAD = \frac{1}{n} \sum |Y_i - \bar{Y}|$

Mean Square Deviation: $MSD = \frac{1}{n} \sum (Y_i - \bar{Y})^2$

Διακύμανση (Variance): $s^2 = \frac{1}{n-1} \sum (Y_i - \bar{Y})^2$

Τυπική Απόκλιση (Standard Deviation) : $s = \sqrt{s^2}$

Συνδιακύμανση (Covariance): $COV_{xy} = \frac{1}{n-1} \sum (X_i - \bar{X})(Y_i - \bar{Y})$

Συσχέτιση (Correlation): $r_{xy} = \frac{COV_{xy}}{s_x s_y}$, $+1 \geq r_{xy} \geq -1$

Αυτοδιακύμανση (Autocovariance): $C_k = \frac{1}{n} \sum_{t=k+1}^n (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})$

Αυτοσυσχέτιση (Autocorrelation lagged k) $r_{kk} = \frac{\sum_{t=k+1}^n (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2}$

Κριτήρια αξιολόγησης των μεθόδων πρόβλεψης

Τα κριτήρια αυτά βασίζονται στις τιμές των αποκλίσεων των προβλεπόμενων τιμών από τις αντίστοιχες πραγματικές τιμές της χρονοσειράς. Όσο πιο μικρές είναι οι αποκλίσεις αυτές, τόσο πιο καλή θεωρείται ότι είναι η μέθοδος πρόβλεψης που χρησιμοποιήθηκε, δεδομένου ότι είναι σε θέση να προσδιορίζει, με σχετικά μεγάλη ακρίβεια, τις τιμές της χρονοσειράς. Αντίθετα, όσο περισσότερο απέχουν οι πραγματικές τιμές από τις προβλεπόμενες, τόσο λιγότερο αξιόλογη είναι η χρησιμοποιούμενη μέθοδος πρόβλεψης.

$$\text{Σφάλμα : } e_t = Act - F$$

$$\text{Μέσο Σφάλμα : } MAE = \frac{1}{n} \sum_{t=1}^n |e_t|$$

(δείκτης bias)

$$\text{Μέσο Απόλυτο Σφάλμα : } ME = \frac{1}{n} \sum_{t=1}^n e_t$$

$$\text{Μέσο Τετραγωνικό : } MSE = \frac{1}{n} \sum_{t=1}^n e_t^2$$

Μέσο Ποσοστιαίο Σφάλμα : $MPE = \frac{1}{n} \sum_{t=1}^n \left(\frac{Act - F}{Act} \right) 100$

Μέσο Απόλυτο Ποσοστιαίο Σφάλμα: $MAPE = \frac{\sum_{t=1}^n \left| \frac{F_t - Act_t}{Act_t} \right|}{n} 100$

Στον παρακάτω πίνακα παρουσιάζονται οι δείκτες ακρίβειας που συναντώνται στη βιβλιογραφία

| Σύντμηση | Πλήρης Ονομασία |
|---------------|--|
| MSE | Mean Squared Error |
| RMSE | Root Mean Squared Error |
| MAE | Mean Absolute Error |
| MdAE | Median Absolute Error |
| MAPE | Mean Absolute Percentage Error |
| MdAPE | Median Absolute Percentage Error |
| sMAPE | Symmetric Mean Absolute Percentage Error |
| sMdAPE | Symmetric Median Absolute Percentage error |
| MRAE | Mean Relative Absolute Error |
| MdAPE | Median Relative Absolute Error |
| GMRAE | Geometric Mean Relative Absolute Error |

| Σύντμηση | Πλήρης Ονομασία |
|----------------|----------------------------------|
| ReMAE | Relative Mean Absolute Error |
| ReRMSE | Relative Root Mean Squared Error |
| LMR | Log Mean Squared Error Ratio |
| PB | Percentage Beter |
| PB(MAE) | Percentage Better (MAE) |
| PB(MSE) | Percentage Better(MSE) |
| MAsE | Mean Absolute scaled Error |
| MdAsE | Median Absolute Scaled Error |

Οι στατιστικοί δείκτες σφαλμάτων διαχωρίζονται σε τρεις κατηγορίες (Hyndman and Koehler ,2006). Δείκτες που εξαρτώνται από την κλίμακα των δεδομένων είναι οι MSE,RMSE,MAE,MdAE των οποίων η τιμή εξαρτάται από την απόλυτη τιμή των δεδομένων.Δείκτες που βασίζονται σε ποσοστιαία σφάλματα είναι οι MAPE , MdAPE, sMAPE και sMdAPE. Δείκτες που βασίζονται σε σχετικά σφάλματα είναι οι MRAE, MdRAE και GMRAE , το βασικό πρόβλημα είναι οι μικρές τιμές για το σφάλμα αναφοράς .

Μέθοδοι εξομάλυνσης

Για τη διενέργεια προβλέψεων υπάρχει ένας μεγάλος αριθμός μεθόδων που μπορούν να χρησιμοποιηθούν ανάλογα με την εξεταζόμενη περίπτωση. Οι περισσότερες από αυτές αναπτύχθηκαν πριν από μερικές δεκαετίες και εφαρμόστηκαν σε ευρεία κλίμακα λόγω της ραγδαίας ανάπτυξης κατάλληλων υπολογιστικών προγραμμάτων. Στόχος μας είναι η χρησιμοποίηση εκείνων των μεθόδων που είναι εύχρηστες, εφαρμόζονται συχνά και έχουν αποδειχθεί στην πράξη ότι δίνουν καλά αποτελέσματα. Μάλιστα, αν ληφθεί υπόψη η ευκολία με την οποία προσδιορίζονται οι προβλέψεις με τη χρήση των H/Y , τότε το κύριο βάρος της έρευνας θα πρέπει να εστιάζεται αφενός στην αξιολόγηση των αποτελεσμάτων και αφετέρου στην επιλογή της καταλληλότερης μεθόδου πρόβλεψης.

Η *ανάλυση χρονοσειρών (time series analysis)* ασχολείται αποκλειστικά με τη διερεύνηση της διαχρονικής συμπεριφοράς των τιμών μιας μεταβλητής, οι παρατηρήσεις της οποίας προέρχονται από χρονοσειρά. Η πρόβλεψη των μελλοντικών τιμών της μεταβλητής σύμφωνα με την γραμμική ανάλυση χρονοσειρών μπορεί να προέλθει από τις ακόλουθες τρεις κατηγορίες μεθόδων προβλέψεων :

- **Μέθοδοι εξομάλυνσης (smoothing methods):** είναι τεχνικές με τις οποίες προσδιορίζονται οι μελλοντικές τιμές μιας μεταβλητής με βάση τον τρόπο εφαρμογής τους. Οι τεχνικές αυτές ονομάζονται μέθοδοι εξομάλυνσης, διότι η δημιουργία των προβλέψεων προέρχεται από την εξομάλυνση της διαχρονικής εξέλιξης των τιμών της μεταβλητής, ώστε να αναγνωριστεί καλύτερα ο τρόπος συμπεριφοράς της. Ορισμένες από αυτές τις μεθόδους μπορούν να εφαρμοστούν και σε περιπτώσεις μικρού αριθμού παρατηρήσεων της μεταβλητής.

- **Διάσπαση Χρονοσειρών (time series decomposition):** στηρίζεται στην υπόθεση ότι οι τιμές μιας χρονοσειράς σχηματίζονται από τα στοιχεία που τη συνθέτουν και τα οποία είναι η *τάση*, η *κυκλικότητα*, η *εποχικότητα* και η *μη-κανονικότητα*. Για τη δημιουργία των προβλέψεων με τη μέθοδο αυτή, η χρονοσειρά διασπάται στα ανωτέρω τέσσερα συνθετικά της στοιχεία και προσδιορίζεται η επιρροή που έχει καθένα από αυτά στη διαμόρφωση των τιμών της μεταβλητής.
- **Ανάλυση ARIMA (Autoregressive Intergrated Moving Average Analysis):** γνωστή και ως Box-Jenkins ανάλυση, είναι μια περισσότερο σύνθετη στατιστική μέθοδος από τις δύο προηγούμενες, με την οποία αναγνωρίζεται ο τρόπος δημιουργίας των τιμών της μεταβλητής. Με τη μέθοδο αυτή επιλέγεται από μια μεγάλη κατηγορία υποδειγμάτων, εκείνο που θεωρείται ως το πιο κατάλληλο για τη διαμόρφωση των προβλέψεων.

Στατιστικές μέθοδοι Πρόβλεψης

Μέθοδος Naive

Η απλοϊκή μέθοδος Naive αποτελεί την απλούστερη στατιστική μέθοδο πρόβλεψης. Η πρόβλεψη για την κάθε χρονική περίοδο t είναι απλά η παρατηρούμενη τιμή της προηγούμενης περιόδου $t-1$, δηλαδή:

$$F_t = Y_{t-1}$$

Η μέθοδος Naive δεν συνίσταται για μακροπρόθεσμες προβλέψεις, αλλά μπορεί κάλλιστα να χρησιμοποιηθεί ως σημείο αναφοράς για άλλες πιο πολύπλοκες μεθόδους.

Να σημειωθεί πως όταν πραγματοποιούνται προβλέψεις σε αποεποχικοποιημένες χρονοσειρές η μέθοδος αυτή ονομάζεται “Naive 2” καθώς δεν πραγματοποιείται στα αρχικά δεδομένα.

Μέθοδοι εκθετικής εξομάλυνσης

Η εκθετική εξομάλυνση είναι μια μέθοδος πρόβλεψης η οποία προεκτείνει στοιχεία του προτύπου των ιστορικών δεδομένων, όπως τάσεις και εποχιακούς κύκλους, στο μέλλον. Οι προβλέψεις υπολογίζονται μετά από εξομάλυνση των δεδομένων, προκειμένου να απομονωθούν τα πραγματικά πρότυπα από τις τυχαίες διακυμάνσεις.

Η δημοτικότητα των μεθόδων αυτών οφείλεται στην απλότητα των μοντέλων που υιοθετούν, τις περιορισμένες απαιτήσεις τους σε αποθήκευση δεδομένων και τον μειωμένο υπολογιστικό φόρτο. Εμπειρικές μελέτες αποδεικνύουν ότι οι μέθοδοι εκθετικής εξομάλυνσης παρουσιάζουν ικανοποιητικά ποσοστά ακρίβειας σε σχέση με πιο πολύπλοκες μεθόδους πρόβλεψης. Το γεγονός αυτό οφείλεται στο ότι οι μέθοδοι εκθετικής εξομάλυνσης δεν επηρεάζονται από τις ιδιομορφίες των προτύπων των δεδομένων ή από περιστασιακά εμφανιζόμενες ακραίες τιμές, οι οποίες παρατηρούνται σε επιχειρησιακά δεδομένα.

Απλή εκθετική εξομάλυνση σταθερού επιπέδου (SES)

Το μοντέλο σταθερού επιπέδου αναφέρεται και ως απλή εκθετική εξομάλυνση (SES) και περιγράφεται μαθηματικά από τις εξισώσεις

$$e_t = Y_t - F_t$$

$$S_t = S_{t-1} + \alpha \cdot e_t$$

$$F_{t+1} = S_t$$

όπου,

t η χρονική περίοδος

Y_t η πραγματική τιμή των δεδομένων F_t η πρόβλεψη τη χρονική στιγμή t

e_t το σφάλμα (απόκλιση πραγματικής τιμής από πρόβλεψη)

S_t το επίπεδο της χρονοσειρές

α ο συντελεστής εξομάλυνσης, λαμβάνει τιμές στο διάστημα $[0, 1]$

Προκειμένου να ξεκινήσει η διαδικασία υπολογισμού του μοντέλου πρόβλεψης, πρέπει να οριστεί ένα αρχικό επίπεδο (S_0). Ως αρχικό επίπεδο, συνήθως, χρησιμοποιείται: ο μέσος όρος όλων των παρατηρήσεων, ο μέσος όρος των n πρώτων παρατηρήσεων, η πρώτη παρατήρηση, το σταθερό επίπεδο από το μοντέλο της απλής γραμμικής παλινδρόμησης

Η επιλογή του βέλτιστου συντελεστή εξομάλυνσης προκύπτει από δύο παράγοντες οι οποίοι αλληλεξαρτώνται. Αφενός όσο περισσότερος θόρυβος υπάρχει στα δεδομένα της χρονοσειράς, τόσο μικρότερη θα πρέπει να είναι η τιμή του συντελεστή εξομάλυνσης προκειμένου να αποφευχθεί η υπερβολική αντίδραση στο θόρυβο, αφετέρου αν ο μέσος όρος μεταβάλλεται, ο

συντελεστής εξομάλυνσης θα πρέπει να είναι μεγάλος ώστε οι προβλέψεις να παρακολουθούν τις μεταβολές που παρουσιάζουν τα δεδομένα. Η γραμμική αναζήτηση αποτελεί την πιο διαδεδομένη μέθοδο υπολογισμού, η οποία ελαχιστοποιώντας το μέσο τετραγωνικό σφάλμα (MSE) και δίνει ως έξοδο την βέλτιστη τιμή του συντελεστή εξομάλυνσης .

Εκθετική εξομάλυνση γραμμικής τάσης (Holt exponential smoothing)

Το μοντέλο γραμμικής τάσης αποτελεί προέκταση του μοντέλου της απλής εκθετικής εξομάλυνσης, στο οποίο μπορεί πλέον να γίνει διαχείριση της συνιστώσας της τάσης που παρατηρείται στα δεδομένα πολλών χρονοσειρών. Οι εξισώσεις που περιγράφουν το μοντέλο είναι οι εξής :

$$e_t = Y_t - F_t$$

$$S_t = S_{t-1} + T_{t-1} + \alpha e_t$$

$$T_t = T_{t-1} + \beta e_t$$

$$F_{t+m} = S_t + mT_t$$

Με e συμβολίζεται το σφάλμα ως η απόκλιση της πραγματικής τιμής από την προβλεπόμενη, με X_t η πραγματική τιμή της χρονοσειράς(την οποία έχουμε πλέον διαθέσιμη), με S το επίπεδο της χρονοσειράς, με T η τάση της χρονοσειράς, με F η προβλεπόμενη τιμή, με α ο συντελεστής

εξομάλυνσης του επιπέδου και με γ ο συντελεστής εξομάλυνσης της τάσης της χρονοσειράς. Τέλος, με t συμβολίζεται η χρονική περίοδος και με m ο χρονικός ορίζοντας της πρόβλεψης.

Η τιμή του αρχικού επιπέδου της χρονοσειράς, S_0 , υπολογίζεται όπως και στη απλή μέθοδο εκθετικής εξομάλυνσης, ενώ ως τιμή της αρχικής τάσης, T_0 , λαμβάνεται ένα από τα παρακάτω:

Η διαφορά της δεύτερης και της πρώτης παρατήρησης, Η διαφορά μιας τυχαίας παρατήρησης της χρονοσειράς με την προηγούμενη ακριβώς παρατήρησή της, διαιρεμένης με την απόσταση των παρατηρήσεων. Η σταθερά της κλίσης από το μοντέλο απλής γραμμικής παλινδρόμησης. Και στη μέθοδο αυτή η επιλογή των αρχικών τιμών των S και T είναι σημαντικό κομμάτι της διαδικασίας που ακολουθείται, εφόσον επηρεάζουν το μοντέλο και τις παραγόμενες προβλέψεις και γίνεται με βάση τα ποιοτικά χαρακτηριστικά της χρονοσειράς.

Οι συντελεστές α και γ λαμβάνουν τιμές στο διάστημα $[0,1]$ και η εύρεση της βέλτιστης τιμής τους στηρίζεται, όπως και στην προηγούμενη μέθοδο, στην ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος. Είναι σύνηθες η βέλτιστη τιμή του συντελεστή εξομάλυνσης τάσης, γ , να είναι μικρότερη από την αντίστοιχη βέλτιστη τιμή του συντελεστή εξομάλυνσης επιπέδου, α . Αυτό συμβαίνει επειδή η τιμή της τάσης κάθε περιόδου είναι συνήθως πολύ μικρότερη από την τιμή του επιπέδου της χρονοσειράς.

Πλέον οι προβλέψεις δεν είναι σταθερές όπως στην απλή εκθετική εξομάλυνση, αλλά εμφανίζουν τάση, όπως φαίνεται και από την τελευταία εξίσωση, της προβλεπόμενης τιμής. Έχει παρατηρηθεί ότι το συγκεκριμένο μοντέλο πρόβλεψης παράγει αρκετές φορές ιδιαίτερα αισιόδοξα αποτελέσματα σε σχέση με τις πραγματικές τιμές που καταγράφονται, ιδίως για μεσοπρόθεσμες και μακροπρόθεσμες προβλέψεις.

Εκθετική εξομάλυνση μη γραμμικής τάσης

Για τις περιπτώσεις στις οποίες το μοντέλο εξομάλυνσης για γραμμική τάση αποτυγχάνει να αποτυπώσει σωστά τη μελλοντική πορεία των χρονοσειρών (εξαιτίας της υπερεκτίμησης που το χαρακτηρίζει), αναπτύχθηκε το μοντέλο μη γραμμικής τάσης (Gardner και McKenzie το 1985), το οποίο προκαλεί βαθμιαία μείωση του μεγέθους αύξησης των τιμών της χρονοσειράς σε κάθε χρονική περίοδο. Οι εξισώσεις που περιγράφουν το μοντέλο εκθετικής εξομάλυνσης μη γραμμικής τάσης είναι οι εξής:

$$e_t = Y_t - F_t$$

$$S_t = S_{t-1} + T_{t-1} + \alpha e_t$$

$$T_t = T_{t-1} + \beta e_t$$

$$F_{t+m} = S_t + \sum_{i=1}^m \varphi^i T_t$$

Οι περισσότεροι συμβολισμοί είναι γνωστοί ήδη από τις προηγούμενες μεθόδους εκθετικής εξομάλυνσης. Συγκεκριμένα, με e συμβολίζεται η απόκλιση της πραγματικής τιμής από την προβλεπόμενη, με X_t η πραγματική τιμή της χρονοσειράς την τρέχουσα περίοδο t , με S το επίπεδο της χρονοσειράς, με T η τάση της χρονοσειράς, με F η προβλεπόμενη τιμή, με α ο συντελεστής εξομάλυνσης του επιπέδου και με γ ο συντελεστής εξομάλυνσης της τάσης της χρονοσειράς. Και, όπως ήδη σημειώθηκε, με t συμβολίζεται η χρονική περίοδος, ενώ με m ο χρονικός ορίζοντας της πρόβλεψης.

Υπεισέρχεται όμως πλέον μία παράμετρος που ελέγχει το ρυθμό αύξησης των τιμών για την αφαίρεση της θετικής προκατάληψης (που παρουσιάζει το μοντέλο γραμμικής τάσης). Η παράμετρος αυτή ονομάζεται παράμετρος διόρθωσης της τάσης, φ , η οποία μπορεί να λάβει

τιμές θετικές(ή και τη μηδενική τιμή) χωρίς άνω περιορισμό, σε αντίθεση με τους συντελεστές α και γ . Είναι, ωστόσο, σημαντική η κατηγοριοποίηση της μορφής του μοντέλου, ανάλογα με τα διαστήματα τιμών της παραμέτρου ϕ :

$\phi = 0$ προκύπτει το μοντέλο απλής εκθετικής εξομάλυνσης, καθώς η τάση T δεν έχει κανένα ρόλο στον υπολογισμό της πρόβλεψης.

$0 < \phi < 1$ προκύπτει το μοντέλο φθίνουσας τάσης (Damped Exponential Smoothing), το οποίο μειώνει σταδιακά την επίδραση της τάσης στο αποτέλεσμα της πρόβλεψης και ελέγχει με τον τρόπο αυτό το ρυθμό αύξησης των τιμών τους.

$\phi = 1$ προκύπτει το μοντέλο γραμμικής τάσης όπως φαίνεται από την τελευταία εξίσωση.

$\phi > 1$ προκύπτει το μοντέλο εκθετικής τάσης, το οποίο είναι υπεραισιόδοξο και οδηγεί σε υψηλές τιμές πρόβλεψης. Έτσι συνίσταται η χρήση του μόνο για πρόβλεψη ζήτησης ενός προϊόντος στα αρχικά στάδια εμφάνισής του στην αγορά.

Για την αποφυγή της παραγωγής προκατειλημμένων προβλέψεων λόγω λανθασμένης βελτιστοποίησης του συντελεστή ϕ και για επίτευξη καλύτερων αποτελεσμάτων, περιορίζεται το διάστημα τιμών της παραμέτρου ϕ στο διάστημα $[0,1]$. Η επιλογή των αρχικών τιμών του επιπέδου και της τάσης της χρονοσειράς συνίσταται να γίνεται με εφαρμογή της γραμμικής παλινδρόμησης και ο υπολογισμός του βέλτιστου συνδυασμού των συντελεστών α , γ και ϕ με τη διαδικασία ελαχιστοποίησης του μέσου τετραγωνικού σφάλματος, όπως και στις προηγούμενες μεθόδους.

Έχει παρατηρηθεί ότι το μοντέλο μη γραμμικής τάσης παράγει πιο ακριβείς προβλέψεις σε σχέση με το μοντέλο γραμμικής τάσης, που είναι όμως πιο συχνά χρησιμοποιούμενο. Μπορεί να πει κανείς ότι το μοντέλο μη γραμμικής τάσης δίνει ικανοποιητικά αποτελέσματα σε περιπτώσεις μεγάλου πλήθους χρονοσειρών, οπότε και απαιτείται πλήρως αυτοματοποιημένο σύστημα στην παραγωγή προβλέψεων και μάλιστα όσο μεγαλύτερος ο χρονικός ορίζοντας πρόβλεψης, τόσο καλύτερο είναι το μοντέλο αυτό (συγκριτικά με τις υπόλοιπες μεθόδους εκθετικής εξομάλυνσης). Τέλος, ικανοποιητικά αποτελέσματα δίνει και σε περιπτώσεις αδυναμίας επιλογής ορισμένου μοντέλου πρόβλεψης.

Μοντέλα παλινδρόμησης (Regression Models)

Η παλινδρόμηση πειλαμβάνει διαδικασίες και τεχνικές μοντελοποίησης και ανάλυσης μεταβλητών . Το αντικείμενο της είναι η εύρεση συσχετίσεων μεταξύ μιας εξαρτημένης και περισσότερων ανεξάρτητων μεταβλητών. Συγκεκριμένα η ανάλυση της παλινδρόμησης μας φανερώνει τον τρόπο που αλλάζει η τιμή της εξαρτημένης μεταβλητής, σε σχέση με την μεταβολή μιας οποιασδήποτε ανεξάρτητης μεταβλητής, ενώ οι υπόλοιπες ανεξάρτητες μεταβλητές παραμένουν σταθερές. Η εφαρμογή της ανάλυσης παλινδρόμησης είναι συχνή για την εκτίμηση της εξαρτημένης μεταβλητής, έχοντας συγκεκριμενοποιήσει τις τιμές των ανεξάρτητων μεταβλητών. Με τον τρόπο αυτό προσδιορίζεται η μέση τιμή της προσδοκώμενης εξαρτημένης μεταβλητής , εφόσον οι ανεξάρτητες διατηρούνται.

Στην ανάλυση της παλινδρόμησης , ενδιαφέρον παρουσιάζει η διακύμανση της εξαρτημένης τιμής γύρω από την εξίσωση της παλινδρόμησης , η οποία περιγράφεται απο πιθανοτική κατανομή . Παρόλο που η παλινδρόμηση χρησιμοποιείται σε θέματα πρόβλεψης , ο βασικός

ρόλος της είναι η ανάλυση και η κατανόηση των σχέσεων μεταξύ ανεξάρτητων και εξαρτημένων μεταβλητών . Με την βοήθεια της παλινδρόμησης μπορούμε να αποφανθούμε ποίες απο τις εξεταζόμενες μεταβλητές είναι συσχετισμένες επαρκώς με την ανεξάρτητη μεταβλητή αλλά και την μορφή της συσχέτισης.

Απλή γραμμική παλινδρόμηση

Υποθέτουμε την ύπαρξη σχέσης ανάμεσα στην μεταβλητή πρόβλεψης - εξαρτημένη μεταβλητή και σε μια άλλη μεταβλητή πρόβλεψης - ανεξάρτητη μεταβλητή επίσης υποθέτουμε οτι η σχέση αυτή είναι γραμμική .

Σκοπός της απλής γραμμικής παλινδρόμησης είναι η έκφραση της σχέσης ανάμεσα στις μεταβλητές X και Y με την εξίσωση μιας ευθείας γραμμής :

$$\hat{Y}_i = a + \beta X_i$$

i η χρονική περίοδος

a δηλώνει την τεταγμένη του σημείου τομής της ευθείας με τον άξονα των εξαρτημένων μεταβλητών (όταν δηλαδή η ανεξάρτητη τεθεί ίση με μηδέν)

b δηλώνει την κλίση της ευθείας , δηλαδή το ρυθμό μεταβολής του Y ανα μονάδα αύξησης του X

X_i οι τιμές της ανεξάρτητης μεταβλητής

Y_i οι τιμές της εξαρτημένης μεταβλητής

\hat{Y}_i οι παραγόμενες από το μοντέλο τιμές της εξαρτημένης μεταβλητής

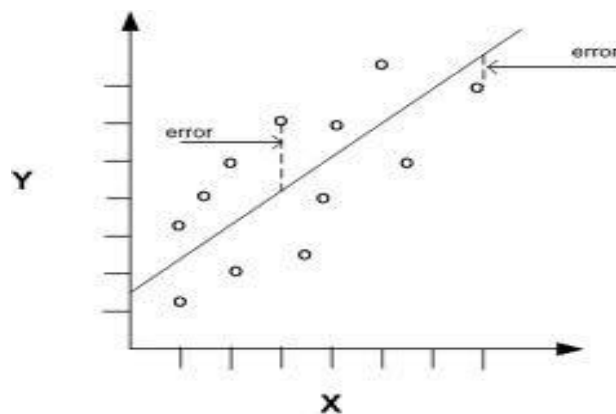
\bar{X} η μέση τιμή των n ανεξάρτητων μεταβλητών

\bar{Y} η μέση τιμή των n εξαρτημένων μεταβλητών

n ο αριθμός των γνωστών παρατηρήσεων

Έχοντας δεδομένες n πραγματικές τιμές (διάνυσμα Y) της εξαρτημένης μεταβλητής , υπολογίζουμε τους συντελεστές a και b ώστε να ελαχιστοποιείται το άθροισμα των τετραγώνων των διαφορών των πραγματικών τιμών Y από τις τιμές που προκύπτουν από την εξίσωση παλινδρόμησης :

$$(a, b) | \min [\sum_{i=1}^n (Y_i - \hat{Y}_i)^2]$$



Η μέθοδος αυτή ονομάζεται ελαχίστων τετραγώνων και η λογική της είναι η ελαχιστοποίηση της απόστασης των πραγματικών παρατηρήσεων Y από την βέλτιστη γραμμή παλινδρόμησης .

$$a = \bar{Y} - b\bar{X}$$

$$b = \frac{\sum_{i=1}^n [(X_i - \bar{X})(Y_i - \bar{Y})]}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

Με \bar{X} και \bar{Y} συμβολίζονται οι μέσες τιμές των μεταβλητών X και Y

αντίστοιχα, με n ο αριθμός των παρατηρήσεων, βάση των οποίων υπολογίζεται η ευθεία παλινδρόμησης και με b συμβολίζεται η εκτίμηση της εξαρτημένης μεταβλητής μέσω της ευθείας παλινδρόμησης. Η πρόβλεψη με χρήση τη μέθοδο της απλής γραμμικής παλινδρόμησης δίνει μια καλή εικόνα της μέσης και της μακροπρόθεσμης συμπεριφοράς του υπό μελέτη μεγέθους. Σε περίπτωση που η σχέση ανάμεσα σε δύο μεταβλητές, την εξαρτημένη και την ανεξάρτητη, δεν είναι γραμμική, μπορεί και πάλι να εφαρμοστεί η μέθοδος της απλής γραμμικής παλινδρόμησης, αφού πρώτα γίνει μετασχηματισμός της σχέσης των δύο μεταβλητών σε γραμμική.

Να σημειωθεί ότι το πραγματικό μοντέλο της απλής γραμμικής παλινδρόμησης δίνεται από τη μαθηματική σχέση:

$$\hat{Y}_i = a + b \cdot X_i + e$$

όπου, e το σφάλμα, δηλαδή η απόκλιση της παρατήρησης από την ευθεία που παριστάνεται από την παραπάνω σχέση.

Συντελεστής γραμμικής συσχέτισης

Για την εφαρμογή της απλής παλινδρόμησης βασική υπόθεση είναι ότι η τιμή της μιας μεταβλητής εξαρτάται από την τιμή ή την μεταβολή της τιμής κάποιας άλλης. Ο συντελεστής γραμμικής συσχέτισης r αποτελεί ένα μέτρο του βαθμού συσχέτισης που μπορεί να υπάρχει μεταξύ δυο μεταβλητών. Ο συντελεστής αυτός παίρνει τιμές από 0 (καμμία συσχέτιση) έως ± 1 (απόλυτη συσχέτιση). Όταν ο συντελεστής γραμμικής συσχέτισης είναι μεγαλύτερος του μηδενός , οι δυο μεταβλητές καλούνται θετικά συσχετισμένες , ενώ όταν είναι μικρότερος του μηδενός , αρνητικά συσχετισμένες . Ο συντελεστής γραμμικής συσχέτισης :

$$r_{xy} = \frac{\sum_{i=1}^n [(X_i - \bar{X})(Y_i - \bar{Y})]}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

$$b = r_{xy} \frac{\sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2}}$$

Ο συντελεστής γραμμικής συσχέτισης ερμηνεύεται ως ένδειξη της κατεύθυνσης της σχέσης ανάμεσα σε δύο μεταβλητές , αν οι τιμές των δυο μεταβλητών αυξομειώνονται συγχρόνως , ή αν αυξάνεται η τιμή της μιας , της άλλης μειώνεται , ή αν οι τιμές των δύο μεταβλητών μεταβάλλονται ανεξάρτητα η μια απο την άλλη . Ερμηνεύεται επίσης ως ένδειξη του βαθμού συσχέτισης , δηλαδή , όσο η τιμή του συντελεστή συσχέτισης απομακρύνεται από το μηδέν , τόσο πιο ισχυρή θεωρείται η συσχέτιση ανάμεσα στις δύο μεταβλητές.

Μοντέλο Πολλαπλής γραμμικής παλινδρόμησης

Όταν απαιτούνται περισσότερες από μία ανεξάρτητες μεταβλητές το μοντέλο της απλής παλινδρόμησης μπορεί να γενικευθεί μέσω της τεχνικής της πολλαπλής παλινδρόμησης ώστε να συμπεριλάβει όλες τις μεταβλητές που επηρεάζουν την τιμή της μεταβλητής πρόβλεψης .Στην πολλαπλή παλινδρόμηση υπάρχει μια εξαρτημένη μεταβλητή της οποίας η τιμή πρέπει να προβλεφθεί βάσει των τιμών δύο ή περισσότερων ανεξάρτητων μεταβλητών .Η γενική μορφή είναι η εξής :

$$Y = b_0 + b_1 X_1 + \dots + b_k X_k + e$$

Η μεταβλητή Y εκφράζει την εξαρτημένη μεταβλητή

X_1, X_2, \dots, X_k εκφράζουν τις ανεξάρτητες μεταβλητές

b_0, \dots, b_k είναι σταθερές παράμετροι

e δηλώνει τον τυχαίο παράγοντα, ο οποίος θεωρείται κανονικά κατανεμημένος γύρω από το μηδέν

Η εξίσωση της πολλαπλής παλινδρόμησης είναι γραμμική ως προς τους συντελεστές. Όταν υπάρχει μια μόνο μεταβλητή X , τότε το σχήμα της συνάρτησης είναι μια ευθεία γραμμή. Αν υπάρχουν περισσότερες από δυο μεταβλητές τότε η Y παριστάνεται σε επίπεδο. Αν υπάρχουν περισσότερες από δύο ανεξάρτητες μεταβλητές, τότε η Y παριστάνεται σε υπερεπίπεδο.

Η μέθοδος της πολλαπλής παλινδρόμησης αποσκοπεί στον υπολογισμό των άγνωστων παραμέτρων του μοντέλου b_0, b_1, \dots, b_k και της διακύμανσης του τυχαίου παράγοντα, δεδομένου ενός συγκεκριμένου συνόλου δεδομένων όπου μπορεί να εφαρμοστεί η μέθοδος των ελαχίστων τετραγώνων.

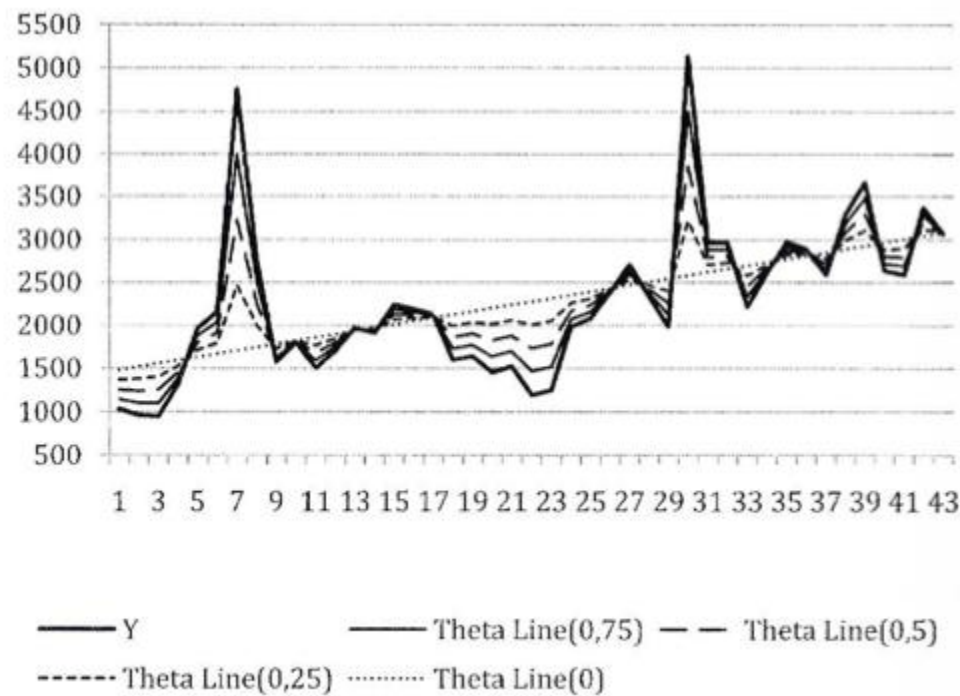
Μοντέλο THETA

Η μέθοδος Theta (Assimakopoulos and Nikolopoulos, 2000) στοχεύει στην τροποποίηση των τοπικών καμπυλοτήτων των χρονοσειρών μέσω της παραμέτρου θ . Αποσυντίθεται η αρχική χρονοσειρά σε δύο ή περισσότερες γραμμές Theta, προεκτείνεται ξεχωριστά κάθε μια και οι προβλέψεις τους συνδιάζονται. Η μέθοδος είναι αρκετά ακριβής για τις μηνιαίες χρονοσειρές και τα μικροοικονομικά δεδομένα.

$$Y^{\theta} = Y_t' \theta$$

$$Y_t'' = Y_t - 2Y_{t-1} + Y_{t-2}$$

Όσο μικρότερη είναι η τιμή της παραμέτρου θ , τόσο μεγαλύτερος είναι ο βαθμός μείωσης των καμπυλοτήτων . Για μηδενική τιμή της παραμέτρου η μέθοδος παράγει γραμμή που ισοδυναμεί με την ευθεία των ελαχίστων τετραγώνων. Να σημειωθεί ότι κάθε γραμμή Theta διατηρεί τη μέση τιμή, αλλά και την κλίση της παλινδρόμησης της αρχικής χρονοσειράς, ενώ δέχονται μεταβολή οι καμπυλότητες και η διακύμανσή της. Η μέθοδος Theta καταλήγει σε καλύτερα αποτελέσματα στις μακροπρόθεσμες προβλέψεις της εξέλιξης των δεδομένων ή ενίσχυση της βραχυπρόθεσμης συμπεριφοράς της χρονοσειράς ανάλογα με την τιμή της παραμέτρου θ .



Σχήμα : Μοντέλο Theta

Αν $\theta=0$ η χρονοσειρά ισοδυναμεί με την ευθεία της απλής γραμμικής παλινδρόμησης.

Αν $\theta = -1$ η χρονοσειρά αντιστοιχεί με τη συμμετρική της αρχικής χρονοσειράς ως προς την ευθεία της γραμμικής παλινδρόμησης.

Αν $\theta > 1$ οι τοπικές καμπυλότητες ενισχύονται και η χρονοσειρά προκύπτει διογκωμένη . Όσο μεγαλύτερος είναι ο βαθμός διόγκωσης , τόσο περισσότερο ενισχύεται η βραχυπρόθεσμη συμπεριφορά της χρονοσειράς

Μοντέλα ARIMA

Τα ολοκληρωμένα αυτοπαλινδρομικά μοντέλα κινητών μέσων όρων (AutoRegressive Integrated Moving Average) είναι στοχαστικά μαθηματικά μοντέλα με τα οποία περιγράφουμε την εξέλιξη κάποιου φυσικου μεγέθους .

Η προσέγγιση των Box-Jenkins στην ανάλυση χρονοσειρών είναι μια μέθοδος εύρεσης ενός στατιστικού υποδείγματος ARIMA που να παριστάνει ικανοποιητικά τη στοχαστική διαδικασία από την οποία προήλθαν τα δεδομένα, δηλαδή το δείγμα μας. Η μέθοδος αυτή εφαρμόζεται όταν η χρονοσειρά δεν είναι στάσιμη και περιλαμβάνει τέσσερα στάδια, την ταυτοποίηση (identification), την εκτίμηση (estimation), και το διαγνωστικό έλεγχο (diagnostic checking) και την πρόβλεψη (forecasting). Τα στοχαστικά μοντέλα εμπεριέχουν τον τυχαίο παράγοντα (τυχαίο σφάλμα) , τις τιμές του μεγέθους που εμφανίστηκαν σε προηγούμενες χρονικές στιγμές και κάποιους άλλους στοχαστικούς παράγοντες . Το μοντέλο που προκύπτει είναι ένας ουσιαστικά γραμμικός συνδυασμός των παραπάνω ποσοτήτων .

Η ταυτοποίηση του υποδείγματος πραγματοποιείται εξετάζοντας τις αυτοσυσχετίσεις και τις μερικές αυτοσυσχετίσεις των ακατέργαστων δεδομένων . Με τον τρόπο αυτό ελέγχουμε αν η χρονοσειρά είναι τυχαία , στάσιμη , αν εμφανίζει κάποιο πρότυπο τάσης ή κάποιο εποχικό πρότυπο . Σε περίπτωση που η χρονοσειρά δεν είναι σταθερή τότε η χρονοσειρά διαφορίζεται

μέχρις ότου να αποκτήσει σταθερότητα και έπειτα δημιουργούνται νέα διαγράμματα αυτοσυσχετίσεων και μερικών αυτοσυσχετίσεων .Η απεικόνιση αυτή αποτυπώνει εάν υπάρχει κάποιο πρότυπο τάσης ή κάποιο εποχικό πρότυπο και παράγεται ένα δοκιμαστικό υπόδειγμα ARIMA . Η εκτίμηση του υποδείγματος περιλαμβάνει την μεταβολή των παραμέτρων ϕ, θ που ορίζουν το υπόδειγμα , οι οποίοι ελαχιστοποιούν το μέσο τετραγωνικό σφάλμα (MSE).Η διαδικασία συνεχίζεται μέχρι οι τελικές τιμές των εκτιμήσεων των παραμέτρων να είναι σημαντικά διαφορετικές από τις αρχικές τιμές της διαδικασίας εκτίμησης.

Η προσέγγιση αυτή έχει έναν αριθμό πλεονεκτημάτων έναντι άλλων μεθόδων ανάλυσης χρονοσειρών. Πρώτον , η μέθοδος είναι λογική και στατιστικά ακριβής . Δεύτερον, αποσπά ένα μεγάλο αριθμό πληροφοριών από τα ιστορικά δεδομένα της χρονοσειράς. Τέλος, επιδρά στην αύξηση της ακριβείας της πρόβλεψης , ενώ συγκρατεί τον αριθμό των παραμέτρων σε ένα ελάχιστο επίπεδο, σε σύγκριση με άλλες παρόμοιες διαδικασίες κατασκευής μοντέλων.

Βασικές έννοιες προχωρημένης ανάλυσης

Η συνάρτηση αυτοσυσχετίσης (autocorrelation function) (ACF)

Ο χρησιμότερος στατιστικός δείκτης στην ανάλυση χρονοσειρών είναι ο συντελεστής αυτοσυσχετίσης.

$$r_k = \frac{\sum_{t=k+1}^n (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^n (Y_{t-k} - \bar{Y})^2}$$

Ο r δηλώνει πως σχετίζονται οι διαδοχικές παρατηρήσεις . Οι συντελεστές αυτοσυσχέτισης για καθυστερήσεις $1, 2, \dots$ δημιουργούν την συνάρτηση αυτοσυσχέτισης .

Μοντέλο "λευκού θορύβου" (white noise model)

Το μοντέλο αυτό ονομάζεται "λευκού θορύβου" και αποτελεί θεμελιώδες μοντέλο σε πολλές τεχνικές ανάλυσης . Η σχέση $Y_t = c + e_t$ περιγράφει ένα απλό τυχαίο μοντέλο . Η Y_t εκφράζεται απο δύο μέρη : c ένα συνολικό επίπεδο και μια συνιστώσα τυχαίου σφάλματος e_t .

Δειγματική κατανομή των αυτοσυσχετίσεων

Για ένα μοντέλο "λευκού θορύβου" η δειγματική θεωρία του r_k είναι γνωστή και συνεπώς μπορούν να μελετηθούν οι ιδιότητες της ACF. Γενικότερα, ένας τρόπος προσέγγισης του προβλήματος είναι να εξετασθεί καθεμιά από τις τιμές του r_k και με βάση το τυπικό σφάλμα (standard error) να ελεγχθεί εάν η τιμή αυτή είναι σημαντικά διάφορη του μηδενός.

Θεωρητικά, όλοι οι συντελεστές αυτοσυσχέτισης μια σειράς τυχαίων αριθμών πρέπει να είναι ίσοι με το μηδέν, αλλά στην πράξη, οι συντελεστές αυτοσυσχέτισης του δείγματος δεν είναι ακριβώς μηδέν επειδή τα δείγματα είναι πεπερασμένα. Έχει δειχθεί, ότι για ένα μοντέλο λευκού θορύβου η κατανομή των συντελεστών αυτοσυσχέτισης μπορεί να προσεγγισθεί από την καμπύλη της κανονικής κατανομής

με μέση τιμή 0 και τυπικό σφάλμα $1/\sqrt{n}$ όπου n ο αριθμός των παρατηρήσεων. Η πληροφορία αυτή μπορεί να χρησιμοποιηθεί για την ανάπτυξη ελέγχων υποθέσεων.

Για παράδειγμα, 95% των συντελεστών αυτοσυσχέτισης του δείγματος πρέπει να ανήκουν στο διάστημα που ορίζεται από τη μέση τιμή συν ή πλην 1.96 standard

errors, δηλαδή για ένα μοντέλο λευκού θορύβου στο $\pm 1.96 / \sqrt{n}$. Εάν αυτό δεν συμβαίνει τότε πιθανώς δεν πρόκειται για μοντέλο λευκού θορύβου.

Στατιστικοί δείκτες Q, Q* (Box-Pierce, Ljung-Box)

Ο δείκτης Q εξετάζει συνολικά τους συντελεστές αυτοσυσχέτισης και ελέγχει εάν το σύνολο είναι σημαντικά διάφορο του μηδενός.

$$Q = n \sum_{k=1}^h r_k^2$$

όπου h η τάξη των μεταβλητών καθυστέρησης και n το πλήθος των παρατηρήσεων.

Ο Q ακολουθεί την κατανομή χ^2 με h-m βαθμούς ελευθερίας όπου m οι παράμετροι του μοντέλου, και η τιμή του συγκρίνεται με τις τιμές του πίνακα χ^2 για να εκτιμηθεί το επίπεδο της σημαντικότητάς του. Για ένα μοντέλο λευκού θορύβου m=0, έτσι σαν εναλλακτικός δείκτης χρησιμοποιείται ο δείκτης Q* ο οποίος επίσης εξετάζει συνολικά τους συντελεστές αυτοσυσχέτισης.

$$Q^* = n(n+2) \sum_{k=1}^h (n-k)^{-1} r_k^2$$

Ο Q* ακολουθεί την ίδια κατανομή με τον Q.

Συντελεστής μερικής αυτοσυσχέτισης (Partial Autocorrelation Coefficient)

Ο συντελεστής μερικής αυτοσυσχέτισης χρησιμοποιείται σαν μέτρο του βαθμού της σχέσης ανάμεσα στην Y_t και την Y_{t-k} , όταν οι επιδράσεις όλων των άλλων μεταβλητών καθυστέρησης 1,2,3 , ... , $k-1$ έχουν αφαιρεθεί. Ο συντελεστής μερικής αυτοσυσχέτισης τάξης k (για καθυστέρηση k), συμβολίζεται με a_k και μπορεί να υπολογισθεί εφαρμόζοντας τη μέθοδο της πολλαπλής γραμμικής παλινδρόμησης με εξαρτημένη μεταβλητή την Y_t και ανεξάρτητες μεταβλητές τις Y_{t-1} , ... , Y_{t-k} :

$$Y_t = b_0 + b_1 Y_{t-1} + \dots + b_k Y_{t-k}$$

Ο συντελεστής a_k ισούται με τον συντελεστή b_k . Πρέπει να σημειωθεί ότι ο πρώτος συντελεστής μερικής αυτοσυσχέτισης a_1 είναι πάντα ίσος με τον πρώτο συντελεστή αυτοσυσχέτισης r_1 .

Έλεγχος της στασιμότητας της χρονοσειράς

Εάν η χρονοσειρά είναι στάσιμη, τότε τα δεδομένα κυμαίνονται γύρω από ένα σταθερό μέσο, ανεξάρτητα του χρόνου, και η διακύμανση παραμένει σταθερή. Συνήθως είναι δυνατόν να ελεγχθεί η στασιμότητα με τη χρήση της γραφικής παράστασης της χρονοσειράς :

- Εάν δεν παρατηρείται αλλαγή της μέσης τιμής κατά μήκος του χρόνου, τότε η χρονοσειρά είναι στάσιμη ως προς τη μέση τιμή.
- Εάν δεν παρατηρείται αλλαγή της διακύμανσης κατά μήκος του χρόνου, τότε η χρονοσειρά είναι στάσιμη ως προς τη διακύμανση.

Το διάγραμμα των αυτοσυσχετίσεων μπορεί επίσης να χρησιμοποιηθεί για τον έλεγχο της στασιμότητας ως προς τη μέση τιμή. Οι αυτοσυσχετίσεις στάσιμων χρονοσειρών φθίνουν στο μηδέν με γρήγορο ρυθμό, ενώ για μη στάσιμες χρονοσειρές φθίνουν με αργό ρυθμό καθώς αυξάνει ο αριθμός των καθυστερήσεων.

Μετατροπή μη στάσιμης χρονοσειράς σε στάσιμη

Τάσεις ή άλλα μη στάσιμα πρότυπα στο επίπεδο της χρονοσειράς, έχουν σαν αποτέλεσμα θετικές αυτοσυσχετίσεις οι οποίες επικρατούν στο διάγραμμα των αυτοσυσχετίσεων. Για το λόγο αυτό, είναι σημαντικό να αφαιρεθούν τα μη στάσιμα πρότυπα ώστε να εμφανισθεί η πραγματική δομή των συσχετίσεων και να αναπτυχθεί το κατάλληλο μοντέλο. Ένας τρόπος να επιτευχθεί αυτό είναι με τη χρήση της μεθόδου της διαφορίσης.

Οι σειρές διαφορών πρώτης τάξης προκύπτουν από τις διαφορές των διαδοχικών παρατηρήσεων της αρχικής χρονοσειράς :

$$Y'_t = Y_t - Y_{t-1}$$

Μερικές φορές η διαφορίση πρώτης τάξης δεν αρκεί για τη μετατροπή μιας μη στάσιμης χρονοσειράς σε στάσιμη και είναι αναγκαίο να διαφοριστούν τα δεδομένα για δεύτερη φορά :

$$Y''_t = Y'_t - Y'_{t-1} = Y_t - 2Y_{t-1} + Y_{t-2}$$

Μοντέλο τυχαίου περιπάτου (random walk)

Ένα μοντέλο μπορεί να γραφεί σαν :

$$Y_t = Y_{t-1} + e_t$$

όπου e_t είναι λευκός θόρυβος.

Το μοντέλο είναι ευρέως γνωστό σαν μοντέλο τυχαίου περιπάτου και χρησιμοποιείται για μη στάσιμα δεδομένα. Τα μοντέλα τυχαίου περιπάτου παρουσιάζουν μεγάλες περιόδους εμφανών τάσεων, οι οποίες μπορούν να αλλάξουν κατεύθυνση με απρόβλεπτο τρόπο.

Εποχιακή διαφορίση

Για μη στάσιμα εποχιακά δεδομένα, η στάσιμη χρονοσειρά προκύπτει με τη χρήση των εποχιακών διαφορών. Σαν εποχιακή διαφορά ορίζεται η διαφορά μιας παρατήρησης και της αντίστοιχης παρατήρησης του προηγούμενου έτους. Έτσι για μηνιαία δεδομένα που παρουσιάζουν εποχικότητα, η εποχιακή διαφορά ορίζεται :

$$Y'_t = Y_t - Y_{12}$$

Μοντέλα χρονοσειρών ARIMA

Τα μοντέλα πολλαπλής γραμμικής παλινδρόμησης περιγράφονται από την εξίσωση :

$$Y = b_0 + b_1 x_1 + \dots + b_p x_p + e$$

όπου Y η εξαρτημένη μεταβλητή (υπό πρόβλεψη μεταβλητή) και X_1, X_2, \dots, X_p οι ανεξάρτητες μεταβλητές (επεξηγηματικές μεταβλητές).

η εξίσωση : $Y_t = c + \varphi_1 Y_{t-1} + e_t$

είναι εξίσωση πολλαπλής παλινδρόμησης, με τη διαφορά ότι οι ανεξάρτητες μεταβλητές είναι προηγούμενες τιμές της υπό πρόβλεψη μεταβλητής ή οι τιμές της μεταβλητής για χρονικές καθυστερήσεις 1, 2, ..., p . Το μοντέλο καλείται αυτοπαλινδρομικό (autoregression ή AR).

Τα αυτοπαλινδρομικά μοντέλα πρέπει να τύχουν διαφορετικού χειρισμού από τα κλασσικά παλινδρομικά μοντέλα για δύο λόγους :

- A. Η βασική υπόθεση της ανεξαρτησίας των σφαλμάτων (υπολοίπων) μπορεί εύκολα να παραβιαστεί στα αυτοπαλινδρομικά μοντέλα, αφού οι επεξηγηματικές μεταβλητές έχουν συνήθως μια εξαρτημένη σχέση καθώς περιγράφουν την εξέλιξη του ίδιου μεγέθους.
- B. Ο τρόπος προσδιορισμού του πλήθους των προηγούμενων τιμών της υπό πρόβλεψη μεταβλητής δεν είναι πάντοτε "ευθύς".

Τα αυτοπαλινδρομικά μοντέλα (AR) μπορούν να ενωθούν αποτελεσματικά με τα μοντέλα κινητού μέσου όρου (MA) και να σχηματίσουν μια χρήσιμη ομάδα μοντέλων χρονοσειρών, τα οποία ονομάζονται αυτοπαλινδρομικά μοντέλα κινητού μέσου όρου (autoregressive moving average ή ARMA models). Τα μοντέλα αυτά μπορούν να χρησιμοποιηθούν μόνο για στάσιμες χρονοσειρές αλλά μπορούν να επεκταθούν και σε μη στάσιμες χρονοσειρές, με τη χρήση της μεθόδου της διαφορίσης. Τότε ονομάζονται ολοκληρωμένα (integrated ή I) αυτοπαλινδρομικά μοντέλα κινητού μέσου όρου (autoregressive integrated moving average ή ARIMA models).

Τα μη εποχιακά ARIMA μοντέλα είναι γνωστά σαν $ARIMA(p, d, q)$ όπου:

AR: p = η τάξη του AR όρου

I: d = η τάξη της διαφορίσης

MA : q = η τάξη του MA όρου

Το μοντέλο λευκού θορύβου ταξινομείται σαν $ARIMA(0,0,0)$, ενώ το μοντέλο τυχαίου περιπάτου σαν $ARIMA(0,1,0)$ [24].

Αυτοπαλινδρομικά μοντέλα πρώτης τάξης

Τα ARIMA(1,0,0) ή AR(1) περιγράφονται από την εξίσωση :

$$Y_t = c + \phi_1 Y_{t-1} + e_t$$

Η παρατήρηση Y_t εξαρτάται από την Y_{t-1} , ενώ ο συντελεστής ϕ_1 παίρνει τιμές ανάμεσα στο -1 και 1. Για $\phi_1=0$ το μοντέλο είναι ισοδύναμο με ένα μοντέλο λευκού θορύβου, ενώ για $\phi_1=1$ το μοντέλο είναι ισοδύναμο με ένα μοντέλο τυχαίου περιπάτου. Γενικά, μπορεί να χρησιμοποιηθεί ένα AR(1) μοντέλο εάν :οι συντελεστές αυτοσυσχέτισης φθίνουν εκθετικά και εάν υπάρχει μόνο ένας στατιστικά σημαντικός συντελεστής μερικής αυτοσυσχέτισης.

Μοντέλα Κινητού Μέσου Όρου πρώτης τάξης

Τα ARIMA(0,0,1) ή MA(1) περιγράφονται από την εξίσωση : $Y_t = c + e_t - \theta_1 e_{t-1}$

Η παρατήρηση Y_t εξαρτάται από τον όρο σφάλματος e_t και το προηγούμενο σφάλμα e_{t-1} , ενώ ο συντελεστής θ_1 παίρνει τιμές ανάμεσα στο -1 και 1. Τα σφάλματα e_t ακολουθούν την κανονική κατανομή με μέση τιμή 0 και διακύμανση 1. Οι συντελεστές μερικής αυτοσυσχέτισης φθίνουν εκθετικά στο μηδέν, ενώ υπάρχει μόνο μία σημαντική αυτοσυσχέτιση.

Διαδικασίες ARMA

Τα βασικά στοιχεία των AR και MA μοντέλων μπορούν να συνδυαστούν για την ανάπτυξη πλήθους μοντέλων. Για παράδειγμα η παρακάτω εξίσωση συνδυάζει ένα AR μοντέλο πρώτης τάξης και ένα MA μοντέλο πρώτης τάξης. Το μοντέλο ονομάζεται ARMA(1,1) ή ARIMA(1,0,1), και χρησιμοποιείται για χρονοσειρές στάσιμες ως προς τη μέση τιμή και την διακύμανση.

$$Y_t = c + \varphi_1 Y_{t-1} + e_t - \theta_1 e_{t-1}$$

Η εξίσωση ενός ARMA μοντέλου με όρους μεγαλύτερης τάξης είναι

$$(1 - \varphi_1 B - \dots - \varphi_p B^p) Y_t = c + (1 - \theta_1 B - \dots - \theta_p B^p) e_t$$

Για τις παραμέτρους ενός ARMA μοντέλου ισχύουν οι περιορισμοί που ισχύουν για τις παραμέτρους των AR και MA μοντέλων.

Διαδικασίες ARIMA

Εάν η χρονοσειρά είναι μη στάσιμη τότε αντί για ένα ARMA μοντέλο, εφαρμόζεται ένα μοντέλο ARIMA(p,d,q). Η πιο απλή περίπτωση είναι το ARIMA(1,1,1) με εξίσωση :

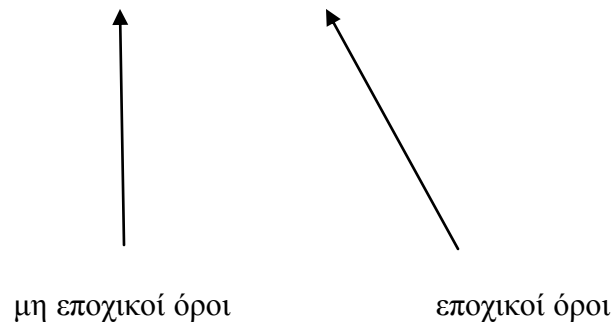
$$(1 - \phi_1 B)(1 - B)Y_t = c + (1 - \theta_1 B)e_t$$

Οι συναρτήσεις αυτοσυσχέτισης (ACF) και μερικής αυτοσυσχέτισης (PACF) των μοντέλων ARIMA(p,d,q) ακολουθούν ένα πλήθος προτύπων που καθιστά αδύνατη τη θέσπιση κανόνων για την αναγνώριση του καταλληλότερου μοντέλου. Η επιλογή όμως ενός κατάλληλου μοντέλου δεν είναι τόσο δύσκολη καθώς διαφορετικά μοντέλα είναι δυνατόν να παράγουν "καλές" προβλέψεις. Στην πράξη σπάνια χρησιμοποιούνται μοντέλα με τιμές p , d , q διαφορετικές από 0, 1 και 2.

Εποχικότητα και μοντέλα ARIMA.

Με τον ίδιο ακριβώς τρόπο που ακόλουθα δεδομένα παρουσιάζουν AR, MA, ARMA και ARIMA ιδιότητες, έτσι και δεδομένα που απέχουν s περιόδους (s η εποχικότητα) είναι δυνατόν να παρουσιάζουν τέτοιες ιδιότητες. Τα μοντέλα ARIMA που έχουν επεκταθεί ώστε να χειρίζονται την εποχικότητα συμβολίζονται ως :

$$\text{ARIMA}(p, d, q)(P, D, Q)_s$$



s συμβολίζεται η εποχικότητα

Εκτίμηση παραμέτρων

Η μέθοδος ελαχίστων τετραγώνων μπορεί να χρησιμοποιηθεί για την εκτίμηση των παραμέτρων ενός ARIMA μοντέλου, αλλά η εφαρμογή της είναι αρκετά δύσκολη όταν στο μοντέλο συμπεριλαμβάνονται MA όροι ($q > 0$). Αντίθετα μπορεί να χρησιμοποιηθεί μια επαναληπτική μέθοδος. Αφού επιλεχθούν αρχικές τιμές για τις παραμέτρους, στη συνέχεια βελτιώνονται μέσω μιας επαναληπτικής διαδικασίας μέχρι να ελαχιστοποιηθεί το άθροισμα των τετραγώνων των σφαλμάτων.

Μια άλλη επαναληπτική μέθοδος η οποία χρησιμοποιείται συχνά είναι η μέθοδος μεγίστης πιθανοφάνειας. Η πιθανοφάνεια ενός συνόλου δεδομένων συμβολίζεται με L και είναι ανάλογη με την πιθανότητα να παραχθούν τα πραγματικά δεδομένα από το μοντέλο. Η μέθοδος μεγίστης πιθανοφάνειας χρησιμοποιείται για την εκτίμηση των τιμών των παραμέτρων που μεγιστοποιούν την πιθανότητα L .

Τα περισσότερα στατιστικά υπολογιστικά πακέτα προσαρμόζουν αυτόματα ένα μοντέλο ARIMA σε μια χρονοσειρά, εκτελούν όλους τους αναγκαίους στατιστικούς ελέγχους και παράγουν εκθέσεις με τις τιμές των παραμέτρων, το τυπικό τους σφάλμα, την t-τιμή τους, το επίπεδο σημαντικότητας, την διακύμανση των υπολοίπων και γενικότερα τις τιμές όλων των στατιστικών δεικτών που χρησιμοποιούνται για τον έλεγχο της καταλληλότητας του μοντέλου

Επιλογή του βέλτιστου μοντέλου ARIMA

Μετά την εκτίμηση των παραμέτρων ενός μοντέλου ARIMA είναι αναγκαία η εκ νέου αναγνώριση προκειμένου να διαπιστωθεί εάν το επιλεγμένο μοντέλο μπορεί να βελτιωθεί. Συγκεκριμένα σε αυτό το στάδιο της διαδικασίας μοντελοποίησης πρέπει:

- A. Εάν προκύψουν συντελεστές στατιστικά μη σημαντικοί, οι αντίστοιχοι όροι πρέπει να αφαιρεθούν από το μοντέλο.
- B. Οι ACF και PACF παρέχουν κάποια καθοδήγηση στην επιλογή ενός απλού AR ή MA μοντέλου. Εάν το καταλληλότερο μοντέλο είναι ένα σύνθετο ARMA μοντέλο, αυτό είναι πολύ δύσκολο να αναγνωρισθεί από τις ACF και PACF. Μετά την επιλογή ενός απλού μοντέλου πρέπει να μελετηθεί εάν αυτό μπορεί να επεκταθεί σε ένα σύνθετο ARMA μοντέλο.
- C. Εάν έχουν εκτιμηθεί περισσότερα από ένα βέλτιστα μοντέλα, πρέπει να εφαρμοσθεί μια μέθοδος επιλογής του καλύτερου από αυτά.

Ένα ικανοποιητικό κριτήριο για την επιλογή του καλύτερου μοντέλου θα μπορούσε να είναι η ελαχιστοποίηση του αθροίσματος ελαχίστων τετραγώνων ή η μεγιστοποίηση της πιθανότητας. Η συγκεκριμένη προσέγγιση όμως δεν δίνει πάντα καλά αποτελέσματα, αφού το άθροισμα ελαχίστων τετραγώνων μπορεί να γίνει μικρότερο και η πιθανότητα μεγαλύτερη, αυξάνοντας απλά τον αριθμό των όρων του μοντέλου.

Ένα ικανοποιητικό κριτήριο είναι το Akaike's Information Criterion ή AIC, το οποίο "τιμωρεί" την εισαγωγή πρόσθετων όρων στο μοντέλο. Αν $m = p + q + P + Q$ είναι το πλήθος των όρων, τότε επιλέγουμε εκείνες τις p, q, P, Q που ελαχιστοποιούν το AIC :

$$AIC = -2 \log L + 2m, \quad L \text{ η πιθανοφάνεια}$$

Διαφορά στις τιμές του AIC μικρότερη του 2 δεν θεωρείται σημαντική και μπορεί να χρησιμοποιηθεί το απλούστερο μοντέλο. Χρησιμοποιούνται επίσης αρκετές παραλλαγές του AIC συμπεριλαμβανομένων επίσης των BIC (Bayesian Information Criterion) και FPE (Final Prediction Error)

Διαγνωστικός έλεγχος

Αν και το επιλεγμένο μοντέλο θεωρείται το καλύτερο από αυτά που εξετάστηκαν, είναι επίσης αναγκαίο να επιβεβαιωθεί η επάρκεια του μοντέλου. Αυτό επιτυγχάνεται με την εξέταση των υπολοίπων (σφαλμάτων), προκειμένου να διαπιστωθεί εάν αυτά ακολουθούν κάποιο πρότυπο.

Τα υπόλοιπα ενός καλού μοντέλου πρόβλεψης πρέπει να είναι "λευκός θόρυβος" και συνεπώς οι ACF και PACF των υπολοίπων δεν πρέπει να παρουσιάζουν στατιστικά σημαντικές αυτοσυσχετίσεις και μερικές αυτοσυσχετίσεις αντίστοιχα. Οι συντελεστές αυτοσυσχέτισης των υπολοίπων μπορούν να εξετασθούν συνολικά με τη χρήση του στατιστικού δείκτη Q^* (Ljung-Box). Εάν η τιμή του Q^* δεν είναι στατιστικά σημαντική τα υπόλοιπα μπορούν να θεωρηθούν μια σειρά "λευκού θορύβου".

Εάν τα υπόλοιπα δεν είναι λευκός θόρυβος τότε το μοντέλο είναι ανεπαρκές και πρέπει να εξετασθούν άλλα μοντέλα ARIMA. Το πρότυπο που ακολουθούν οι στατιστικά σημαντικοί συντελεστές αυτοσυσχέτισης και μερικής αυτοσυσχέτισης των υπολοίπων, υποδεικνύουν τον τρόπο βελτίωσης του μοντέλου. Για παράδειγμα σημαντικές τιμές για εποχιακές καθυστερήσεις υποδεικνύουν την προσθήκη μιας εποχιακής συνιστώσας ή σημαντικές τιμές για μικρές καθυστερήσεις υποδεικνύουν την αύξηση των μη εποχιακών AR ή MA συνιστωσών του μοντέλου.

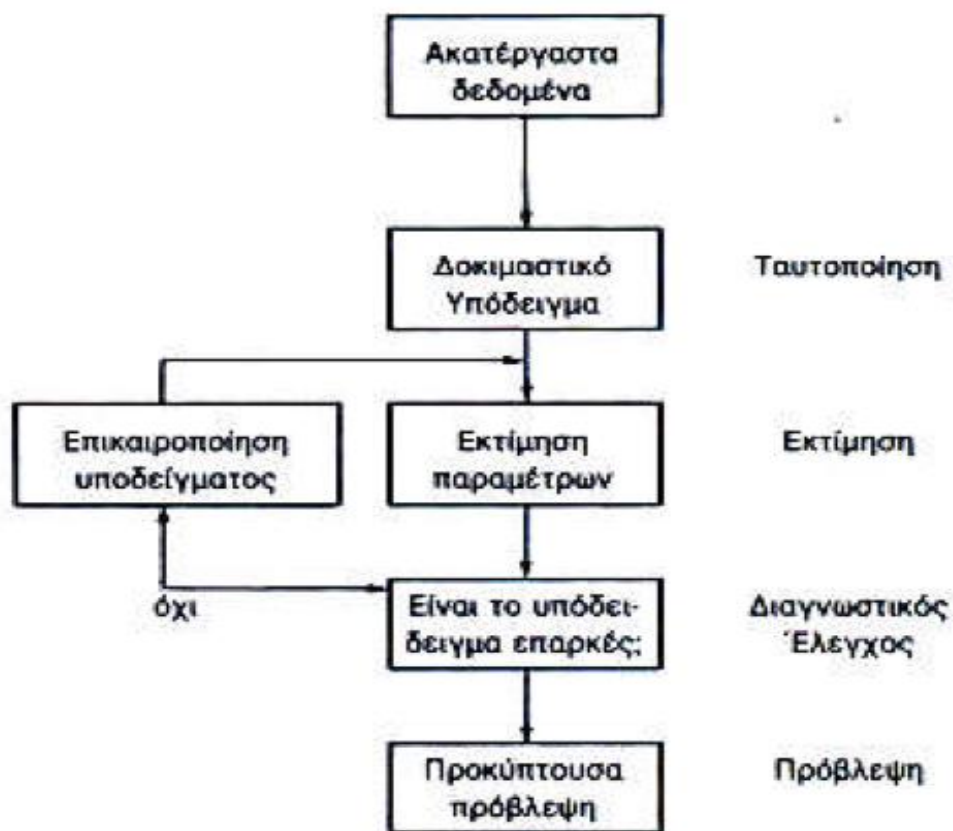
Συνήθως τα μοντέλα με τις μικρότερες AIC τιμές έχουν υπόλοιπα λευκού θορύβου. Περιστασιακά είναι δυνατόν να υιοθετηθεί όχι το μοντέλο με την μικρότερη AIC τιμή αλλά αυτό με τα "καλύτερα" υπόλοιπα .

Διαφόριση και Πρόβλεψη

Η μέθοδος της διαφόρισης έχει μεγάλη επίδραση στις προβλέψεις. Μοντέλα με διαφόριση διαφορετικής τάξης μπορεί να προσαρμόζονται το ίδιο καλά στα ιστορικά δεδομένα αλλά είναι δυνατόν να παράγουν προβλέψεις που διαφέρουν σημαντικά.

Ακολουθώς περιγράφονται συνοπτικά τα χαρακτηριστικά των προβλέψεων που παράγουν τα διάφορα μοντέλα ARIMA (p, d, q) (P, D, Q)_s, καθώς αυξάνει ο ορίζοντας πρόβλεψης.

- $d=0$. Οι προβλέψεις θα συγκλίνουν στη μέση τιμή των ιστορικών δεδομένων εάν συμπεριλαμβάνεται σταθερός όρος στο μοντέλο, διαφορετικά θα συγκλίνουν στο μηδέν. Η διακύμανση των προβλέψεων θα συγκλίνει στη διακύμανση των ιστορικών δεδομένων και συνεπώς τα διαστήματα εμπιστοσύνης επίσης θα συγκλίνουν.
- $d=1$. Οι προβλέψεις θα συγκλίνουν στη τιμή της τελευταίας παρατήρησης εάν δεν συμπεριλαμβάνεται σταθερός όρος στο μοντέλο. Εάν συμπεριλαμβάνεται σταθερός όρος οι μακροπρόθεσμες προβλέψεις θα ακολουθούν μια γραμμική τάση, με κλίση ίση με την τιμή του σταθερού όρου. Και στις δύο περιπτώσεις η διακύμανση των προβλέψεων θα αυξάνει με το χρονικό ορίζοντα και τα διαστήματα εμπιστοσύνης θα αποκλίνουν.
- $d=2$. Εάν δεν συμπεριλαμβάνεται σταθερός όρος στο μοντέλο, οι προβλέψεις θα ακολουθούν μια γραμμική τάση, προεκτείνοντας την τάση που ακολουθούν οι τελευταίες παρατηρήσεις της χρονοσειράς. Εάν συμπεριλαμβάνεται σταθερός όρος οι μακροπρόθεσμες προβλέψεις θα ακολουθούν μια τετραγωνική τάση. Και στις δύο περιπτώσεις η διακύμανση των προβλέψεων θα αποκλίνει και τα διαστήματα εμπιστοσύνης επίσης θα αποκλίνουν.
- Εποχιακή διαφόριση. Οι επιδράσεις της εποχιακής διαφόρισης είναι παρόμοιες. Ειδικότερα, εάν έχει εφαρμοσθεί εποχιακή διαφόριση και στη συνέχεια διαφόριση πρώτης τάξης, οι προβλέψεις θα συμπεριφέρονται όμοια με αυτές του μοντέλου διαφόρισης δεύτερης τάξης ($d=2$), με τη διαφορά ότι θα παρουσιάζουν εποχικότητα



Μέθοδοι διακοπτόμενης Ζήτησης

Μέθοδος CROSTON

Το 1972 ο Croston προτείνει μια εναλλακτική μέθοδο η οποία εξετάζει το μέγεθος της ζήτησης και τον χρόνο μεταξύ των ζητήσεων, εξάγει προβλέψεις εφαρμόζοντας ανεξάρτητα απλή εκθετική εξομάλυνση στις μη μηδενικές τιμές των χρονοσειρών και στα χρονικά διαστήματα

μεταξύ μη μηδενικών τιμών των χρονοσειρών .Στην πράξη διαχωρίζεται η κάθε χρονοσειρά σε δύο επιμέρους , η μια αποτελείται από τα χρονικά διαστήματα μεταξύ μη μηδενικών ζητήσεων (intervals) και η άλλη από το πλήθος των ανεξάρτητων μη μηδενικών ζητήσεων (demands) .Επειτα οι χρονοσειρές προεκτείνονται ανεξάρτητα με την μέθοδο εκθετικής εξομάλυνσης σταθερού επιπέδου . Ως παράμετρος εξομάλυνσης στην εφαρμογή της μεθόδου είναι η χρήση της τιμής $\alpha=0,05$. . Ο υπολογισμός της πρόβλεψης Croston παρουσιάζεται μέσω της εύρεσης του πηλίκου των δυο ανεξάρτητων προβλέψεων

$$F_{\text{croston}} = \frac{F_{\text{demands}}}{F_{\text{intervals}}}$$

Η μέθοδος μπορεί να χρησιμοποιηθεί και για δεδομένα συνεχούς ζήτησης , στην περίπτωση δεδομένων διακοπτόμενης ζήτησης αποτελεί παραλλαγή της απλής εκθετικής εξομάλυνσης. Η μέθοδος Croston ανάγεται σε ένα μεθολογικό πλαίσιο που περιλαμβάνει την αποσύνθεση της αρχικής χρονοσειράς διακοπτόμενης ζήτησης σε δύο σειρές , η μια περιλαμβάνει μη μηδενικές τιμές και η άλλη τα μεσοδιαστήματα μεταξύ αυτών.Την προέκταση κάθε χρονοσειράς που προέκυψαν από την αποσύνθεση με την κατάλληλη μέθοδο πρόβλεψης και τέλος τον συνδυασμό των παραχθέντων προβλέψεων μέσω διαίρεσης της σειράς πρόβλεψης ζήτησης και της σειράς πρόβλεψης μεσοδιαστημάτων προκειμένου να υπολογισθει η τελική σειρά προβλέψεων.

Μέθοδος Syntetos & Boylan Approximation (SBA)

Η μέθοδος Croston παρά τα σημαντικά οφέλη που παρουσιάζει και την ισχυρή ικανότητα πρόβλεψης για δεδομένα διακοπτόμενης φύσης, παρουσιάζει αισιόδοξη τάση στα αποτελέσματα των προβλέψεων της, δηλαδή είναι θετικά προκατειλημμένη (positively biased). Η τροποποιημένη μέθοδος SBA παρουσιάζει μικρότερη προκατάληψη όταν παρατηρούνται πολλές μηδενικές τιμές ζήτησης.

Η μέθοδος SBA αποτελεί τροποποίηση της μεθόδου CROSTON

$$F_{SBA} = \left(1 - \frac{\alpha}{2}\right) \frac{F_{demands}}{F_{intervals}}$$

ΕΠΙΛΟΓΗ ΤΗΣ ΚΑΤΑΛΛΗΛΗΣ ΜΕΘΟΔΟΥ ΠΡΟΒΛΕΨΗΣ

Η επιλογή της κατάλληλης μεθόδου, η οποία θα χρησιμοποιηθεί για την παραγωγή προβλέψεων, δεν είναι μια εύκολη διαδικασία που μπορεί να γίνει από απλούς χρήστες ενός προγράμματος. Για την επιλογή της βέλτιστης μεθόδου πρόβλεψης μπορούμε να εξετάσουμε ορισμένους παράγοντες που επηρεάζουν το βαθμό εφαρμογής μιας μεθόδου και ως εκ τούτου τα αποτελέσματά της. Οι κυριότεροι λοιπόν παράγοντες είναι:

Χρονικός ορίζοντας

Βασικό κριτήριο επιλογής μιας μεθόδου πρόβλεψης αποτελεί το χρονικό διάστημα στο μέλλον στο οποίο θα αναφέρεται η πρόβλεψη. Οι ποιοτικές μέθοδοι χρησιμοποιούνται περισσότερο για μακροπρόθεσμες προβλέψεις ενώ οι ποσοτικές μέθοδοι χρησιμοποιούνται περισσότερο για μεσοπρόθεσμες και βραχυπρόθεσμες προβλέψεις. Επίσης σημαντικό στοιχείο είναι και το πλήθος των περιόδων για το οποίο απαιτείται πρόβλεψη. Ορισμένες τεχνικές είναι κατάλληλες για προβλέψεις που αντιστοιχούν σε 1 ή 2 περιόδους μετά από την πιο πρόσφατη παρατήρηση, ενώ άλλες σε περισσότερες.

Πρότυπο συμπεριφοράς των δεδομένων.

Βασική προϋπόθεση στην πλειοψηφία των μεθόδων πρόβλεψης είναι η αναγνώριση του προτύπου συμπεριφοράς των δεδομένων πάνω στο οποίο θα στηριχθεί η πρόβλεψη. Τα τέσσερα βασικά πρότυπα συμπεριφοράς που συχνά εμφανίζονται στις χρονοσειρές και τις περισσότερες φορές συνυπάρχουν είναι το σταθερό πρότυπο, το πρότυπο της τάσης, το εποχιακό και το

κυκλικό πρότυπο . Είναι σημαντικό , λόγω του ότι η ικανότητα των διαφόρων μεθόδων να παράγουν αξιόπιστες προβλέψεις για διαφορετικά πρότυπα δεδομένων ποικίλλει , η μέθοδος που θα επιλεγεί να είναι κατάλληλη για το συγκεκριμένο πρότυπο.

Κόστος

Το κόστος μιας μεθόδου πρόβλεψης καθορίζεται από τον όγκο των δεδομένων που απαιτεί η μέθοδος και από την πολυπλοκότητα της εφαρμογής της.

Αξιοπιστία

Η αξιοπιστία είναι στενά συνδεδεμένη με το επίπεδο λεπτομέρειας που απαιτείται σε μια πρόβλεψη. Σε ορισμένες περιπτώσεις ένα ποσοστό ακρίβειας $\pm 10\%$ θεωρείται ικανοποιητικό, ενώ σε άλλες έστω και 1 διακύμανση της τάξης του $\pm 5\%$ μπορεί να αποδειχτεί καταστροφική.

Απλότητα και ευκολία εφαρμογής

Έχει αποδειχτεί στην πράξη ότι προτιμώνται μέθοδοι που είναι κατανοητές και εύκολες στην εφαρμογή τους.

ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

Οι χρονοσειρές είναι μια ειδική μορφή στοιχείων όπου οι προηγούμενες τιμές στη σειρά μπορούν να επηρεάσουν μελλοντικές τιμές , εξαρτώμενες από τηπαρουσία υποκείμενων αιτιοκρατικών δυνάμεων (τάσεις , κύκλους ,μη στατιστική συμπεριφορά κ.λπ). Τα νευρωνικά δίκτυα είναι μη γραμμικά μοντέλα που μπορούν να εκπαιδευθούν για να απεικονίσουν τις παρελθοντικές και μελλοντικές τιμές μιας χρονοσειράς και μέσω αυτής της διαδικασίας εξάγουν τις σχέσεις που διέπουν τα δεδομένα . Το νευρωνικό δίκτυο μπορεί να ορισθεί ως τεχνική επεξεργασίας δεδομένων που συσχετίζει κάποιο τύπο ρεύματος εισαγωγής πληροφοριών με ένα ρεύμα εξαγωγής δεδομένων.

Οι στόχοι των νευρωνικών δικτύων μπορούν να διαιρεθούν σε τέσσερις τύπους εφαρμογών :

1. Κατάταξη : Απόφαση σε ποια κατηγορία περιέχεται ένα σχέδιο εισαγωγής
2. Συσχέτιση : Ενεργεί ως προσποελάσιμη μνήμη που ανακαλεί ένα δεδομένο εξόδου δίνοντας κάποιο μέρος ως είσοδο .
3. Κωδικοποίηση: Συμπιέζει ένα δεδομένο εισόδου με την παραγωγή ενός δεδομένου εξόδου με μειωμένη διάσταση.
4. Προσομοίωση : Δημιουργία ενός δεδομένου εξόδου για ένα δεδομένο εισόδου που λειτουργεί ως ερέθισμα .

Πλεονεκτήματα των νευρωνικών δικτύων

Γενίκευση : Τα νευρωνικά δίκτυα καθιερώνονται μέσω μιας φάσης ``κατάρτισης`` , εισάγοντας παραδείγματα και το δίκτυο εκπαιδεύεται για να εξάγει τις σχετικές πληροφορίες .

Ευελιξία : Η σειρά των στόχων στην οποία το νευρωνικό δίκτυο μπορεί να εφαρμοστεί , υπερβαίνει κατά πολύ οποιαδήποτε τεχνική

Μη γραμμική ``διαμόρφωση`` : Διαδικασία χαρτογράφησης δικτύων περιλαμβάνει μη γραμμικές λειτουργίες που καλύπτουν την πολυπλοκότητα του προβλήματος .

Τα νευρωνικά δίκτυα μπορούν να διαιρεθούν σε τρεις κατηγορίες , βάσει της τεχνικής που χρησιμοποιείται για να εκπαιδεύσει επαναληπτικά τις ελεύθερες παραμέτρους (βάρη) στο δίκτυο.

- Εποπτευόμενη μάθηση : Οι στόχοι των δεδομένων εξόδου των δικτύων είναι γνωστοί κατά τη διάρκεια της κατάρτισης έτσι ώστε η διαφορά μεταξύ του επιθυμητού στόχου και της πραγματικής εξόδου , μπορεί να ανατροφοδοτηθεί στο δίκτυο για βελτίωση της απόδοσης
- Ενισχυόμενη Εκμάθηση : Οι στόχοι για κάθε ένα από τα παραδείγματα που εισάγονται στο δίκτυο δεν είναι γνωστοί . Καμία πληροφορία σχετικά με την πραγματική –σωστή έξοδο δεν είναι διαθέσιμη για την κατάρτιση . Σε μερικές περιπτώσεις μόνο η κρίση σχετικά με την εκτίμηση είναι διαθέσιμη , αφού έχει παρουσιαστεί μια ακολουθία εισαγωγών , η οποία δημιουργεί έναν περαιτέρω χωρισμό μεταξύ της κάθε εισαγωγής και της επιθυμητής απόδοσης των δικτύων . Σε αυτή την περίπτωση το ΝΔ τροφοδοτείται και πάλι με δειγματικά πρότυπα εισόδου αλλά δεν τροφοδοτείται με τις επιθυμητές αποκρίσεις σε αυτές τις εισόδους . Σε αυτή την περίπτωση χρησιμοποιείται ένα συνολικό μέτρο της επάρκειας της προκύπτουσας απόκρισης το οποίο μπορεί να οδηγήσει το νευρωνικό δίκτυο στην επιθυμητή συμπεριφορά . Το μέτρο αυτό είναι γνωστό ως ενισχυτικό σήμα (reinforcement signal) και ανατροφοδοτείται στο Ν ευρωνικό δίκτυο

έτσι ώστε να επιβραβεύσει τις ορθές συμπεριφορές και να τιμωρήσει τις λανθασμένες . Η ενισχυτική μάθηση διακρίνεται σε *συσχετιστική* και *μη συσχετιστική* ενισχυτική μάθηση. Στην πρώτη περίπτωση το περιβάλλον τροφοδοτεί πέρα από το ενισχυτικό σήμα και άλλες μορφές πληροφορίας από τις οποίες το ΝΔ πρέπει να αποτυπώσει μια απεικόνιση συσχέτισης αιτίου -αποτελέσματος. Στην δεύτερη περίπτωση η μόνη πληροφορία που δίδεται από το περιβάλλον είναι το ενισχυτικό σήμα και ο προορισμός του ΝΔ είναι να επιλέξει μια μοναδική βέλτιστη ενέργεια.

- Εκμάθηση χωρίς επίβλεψη : Υπάρχει συνολική έλλειψη ανατροφοδότησης σχετικά με τρόπο επιλογής της εκμάθησης . Οι τύποι δικτύων είναι αυτό-οργανωμένοι , χρησιμοποιώντας σχέσεις μεταξύ των προτύπων εισαγωγής . Σε αυτό τον τύπο μάθησης που καλείται αυτό- οργανούμενη μάθηση δεν χρησιμοποιείται εξωτερικός δάσκαλος ούτε μια βάση γνώσης για να επιβλέψει την εκπαίδευση του ΝΔ , το μόνο στοιχείο που μπορεί να χρησιμοποιηθεί για την υλοποίηση της εκπαίδευσης είναι τα διανύσματα εισόδου . Ένα σύστημα μη επιβλεπόμενης μάθησης εξελίσσεται με τέτοιο τρόπο ώστε να εξάγει χαρακτηριστικά ή κανονικότητες από τα παρουσιαζόμενα πρότυπα , χωρίς ωστόσο να έχει την πληροφορία για το ποιες έξοδοι ή ποιες κατηγορίες συσχετίζονται με τα χαρακτηριστικά εισόδου . Με άλλα λόγια το σύστημα μάθησης εντοπίζει ή κατηγοριοποιεί τα διανύσματα εισόδου χωρίς καμία εκ των προτέρων πληροφόρηση από το περιβάλλον. Εξ' αιτίας αυτών η μη επιβλεπόμενη μάθηση συχνά χρησιμοποιείται σε προβλήματα ομαδοποίησης , εξαγωγής εσωτερικών χαρακτηριστικών και ανίχνευσης συμμετριών.

Τα νευρωνικά δίκτυα μη επιβλεπόμενης μάθησης εκπαιδεύονται έτσι ώστε να αποκρίνονται σε διαφορετικά διανύσματα εισόδου με διαφορετικά τμήματα του δικτύου. Το ΝΔ εκπαιδεύεται με τέτοιο τρόπο ώστε να αυξάνει την πυροδότηση του ως απόκριση σε συχνά εμφανιζόμενες εισόδους, γι' αυτό και συχνά ονομάζεται εκτιμητής πιθανοτήτων (probability estimator). Κατά αυτό τον τρόπο το νευρωνικό δίκτυο αναπτύσσει συγκεκριμένες εσωτερικές αναπαραστάσεις οι οποίες κωδικοποιούν τα διάφορα διανύσματα εισόδου.

Τεχνητή νοημοσύνη

Η γεφύρωση του χάσματος ανάμεσα στον άνθρωπο και στις μηχανές αποτελεί το μεγαλύτερο στόχο της Επιστήμης των Υπολογιστών. Όμως, ακόμη και σήμερα, αν και η επιστήμη αυτή έχει εμφανίσει αλματώδη ανάπτυξη, δεν έχει καταφέρει να παρουσιάσει ηλεκτρονικούς υπολογιστές απαλλαγμένους από τα κλασσικά μειονεκτήματα μιας μηχανής. Έτσι, εκτός από ελάχιστες εξαιρέσεις, δεν υπάρχουν σήμερα μηχανές που να είναι σε θέση να επικοινωνήσουν με τον άνθρωπο σε φυσική γλώσσα, να απαντούν σε ερωτήσεις για διάφορα συγκεκριμένα προβλήματα (εκτός εάν είναι εφοδιασμένοι με μια κατάλληλη μέθοδο επίλυσης) ή να αποκτούν εμπειρίες και να μαθαίνουν από τις αποτυχίες και τα λάθη τους. Η περιοχή έρευνας που ασχολείται με αυτού του είδους τα προβλήματα καλείται Τεχνητή Νοημοσύνη.

Τεχνητή Νοημοσύνη ονομάζεται η μελέτη των τεχνικών και των διεργασιών που δίνουν σε έναν υπολογιστή τη δυνατότητα να αποκτά διανοητικές ικανότητες, ανάλογες με αυτές που διαθέτει ο ανθρώπινος εγκέφαλος. Η Τεχνητή Νοημοσύνη έχει ως κύριο στόχο της να γίνει ο υπολογιστής πιο έξυπνος και κατ' επέκταση πιο χρήσιμος, αφού θα είναι σε θέση να ανταποκρίνεται πολύ καλύτερα στις ανάγκες και τις επιθυμίες του ανθρώπου.

Σήμερα ο όρος Τεχνητή Νοημοσύνη χρησιμοποιείται τόσο για τη μελέτη της ανθρώπινης νοημοσύνης, όσο και για την βελτίωση των δυνατοτήτων των υπολογιστών. Οι κυριότερες εφαρμογές της είναι οι εξής:

- Ο προγραμματισμός μιας ακολουθίας ενεργειών για την επίτευξη ενός στόχου (planning).
- Η εξαγωγή συμπερασμάτων (inference) μέσα από αλληλοσυσχετιζόμενα γεγονότα και η λήψη αποφάσεων (decision making) ή και πρόβλεψη (forecasting).
- Η παροχή συμβουλών και συμπερασμάτων μέσα από σύνθετες δομές κανόνων και γεγονότων (expert systems).
- Η εκπαίδευση των υπολογιστών για επικοινωνία με τους ανθρώπους μέσω φυσικών γλωσσών. Αυτό περιλαμβάνει μια ποικιλία εφαρμογών όπως αναγνώριση φωνής, παραγωγή φωνής, κατανόηση κειμένου, κ.τ.λ.
- Η αυτόνομη κίνηση των υπολογιστών και η μετακίνηση από αυτούς αντικειμένων μέσα στο χώρο (robotics).
- Η αναγνώριση αντικειμένων μέσω κάμερας (vision).

Περιοχές Έρευνας της Τεχνητής Νοημοσύνης

Οι περιοχές έρευνας της Τεχνητής Νοημοσύνης που συγκεντρώνουν το μεγαλύτερο ενδιαφέρον σήμερα είναι οι εξής:

1. Έμπειρα Συστήματα (Expert Systems)
2. Γενετικοί Αλγόριθμοι (Genetic Algorithms)
3. Ασαφή Συστήματα (Fuzzy Systems)
4. Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks)

ΕΜΠΕΙΡΑ ΣΥΣΤΗΜΑΤΑ

Τα Έμπειρα Συστήματα αποσκοπούν στην πραγματοποίηση συστημάτων υπολογιστών με δυνατότητες αυτόματης επεξεργασίας των δεδομένων ενός προβλήματος και την εξαγωγή συγκεκριμένων συμπερασμάτων. Χρησιμοποιούν προγραμματιστικές μεθόδους της Τ.Ν. σε συνδυασμό με τις γνώσεις εμπειρογνομόνων σε ειδικά θέματα. Ένα Έμπειρο Σύστημα αποτελείται από τη βάση γνώσης (knowledge base) και το μηχανισμό συμπεράσματος (inference engine), ο οποίος μπορεί να έχει διάφορες μορφές, ανάλογα με την εφαρμογή.

Τα τελευταία τριάντα χρόνια, έχει παρατηρηθεί ένα συνεχώς αυξανόμενο ενδιαφέρον για ανάπτυξη συστημάτων επίλυσης προβλημάτων βασισμένων στις αρχές της Γενετικής Εξέλιξης και της Κληρονομικότητας. Τα μειονεκτήματα των κλασικών μεθόδων αναζήτησης και βελτιστοποίησης, καθώς και η συνεχώς αυξανόμενη ανάγκη για παραγωγή λογισμικού που να μπορεί να εκμεταλλεύεται πιο αποδοτικά τις τεράστιες δυνατότητες του υλικού, ήταν η βασική αιτία που ώθησε τους επιστήμονες σ' αυτήν την αναζήτηση. Αυτού του είδους τα συστήματα λειτουργούν διατηρώντας ένα πληθυσμό κωδικοποιημένων πιθανών λύσεων και εφαρμόζοντας πάνω σε αυτό διάφορες διαδικασίες επιλογής του καλύτερου, καθώς και διάφορους γενετικούς τελεστές. Οι τελεστές αυτοί αντιγράφουν τον τρόπο με τον οποίο αναπαράγονται και μεταλλάσσονται τα χρωμοσώματα των κυττάρων των ζωντανών οργανισμών. Έτσι, περνώντας από γενιά σε γενιά, τα συστήματα αυτά δημιουργούν συνεχώς νέους πληθυσμούς πιθανών λύσεων χρησιμοποιώντας, τόσο κομμάτια και στοιχεία από την προηγούμενη γενιά, όσο και εντελώς καινούρια κομμάτια που δοκιμάζονται για τυχόν καλή απόδοσή τους. Με αυτόν τον τρόπο αξιοποιούν τις πληροφορίες που τους παρέχει το περιβάλλον τους. Επανειλημμένες δοκιμές και πειράματα έχουν δείξει ότι μια "φυσική" αναπαράσταση των πιθανών λύσεων για ένα δεδομένο πρόβλημα σε συνδυασμό με την εφαρμογή σε αυτή μιας οικογένειας γενετικών τελεστών, αποτελεί πολύ χρήσιμο εργαλείο στην προσπάθεια προσέγγισης των πραγματικών λύσεων σε μια πολύ μεγάλη ποικιλία προβλημάτων και εφαρμογών. Αυτό το γεγονός μετατρέπει αυτή την προσέγγιση "φυσικού μοντέλου" σε μια πολλά υποσχόμενη κατεύθυνση όσον αφορά την επίλυση προβλημάτων γενικότερα.

ΓΕΝΕΤΙΚΟΙ ΑΛΓΟΡΙΘΜΟΙ

Το 1958 ο Friedberg , προσπάθησε να προσομοιώσει πολύπλοκα βιολογικά συστήματα με την χρήση ηλεκτρονικού υπολογιστή , ωστόσο τα προγράμματα που προέκυψαν δεν μπορούσαν να εκτελεστούν

Η συστηματική ανάπτυξη όμως προέκυψε από τον John Holland χρησιμοποιώντας σειρές bits για να αναπαραστήσει εγκυρές λειτουργίες. Η βασική συλλήψη προήλθε από τους βιολογικούς μηχανισμούς μίμησης της φύσης.

Οι Γεντικοί Αλγορίθμοι χρησιμοποιούν ορολογία από τον χώρο της Γενετικής .

Αναφέρονται σε άτομα (individuals) μέσα σε ένα πληθυσμό , όπου κάθε άτομο αποτελείται από χρωμοσώματα (chromosomes) . Τα χρωμοσώματα αποτελούνται από γονίδια (genes) που είναι διατεταγμένα σε γραμμική ακολουθία . Κάθε γονίδιο επηρεάζει την κληρονομικότητα ενός ή περισσότερων χαρακτηριστικών . Τα ιδιαίτερα χαρακτηριστικά ενός ατόμου συναντώνται σε συγκεκριμένες θέσεις στο χρωμοσώμα και ονομάζονται loci . Οι διαφορετικές καταστάσεις που μπορεί να πάρει το γονίδιο καλούνται τιμές χαρακτηριστικού γνωρίσματος (alleles) . Περνώντας από γενιά σε γενιά, τα συστήματα αυτά δημιουργούν συνεχώς νέους πληθυσμούς πιθανών λύσεων χρησιμοποιώντας, τόσο κομμάτια και στοιχεία από την προηγούμενη γενιά, όσο και εντελώς καινούρια κομμάτια που δοκιμάζονται για τυχόν καλή απόδοσή τους. Με αυτόν τον τρόπο αξιοποιούν τις πληροφορίες που τους παρέχει το περιβάλλον τους. Επανειλημμένες δοκιμές και πειράματα έχουν δείξει ότι μια "φυσική" αναπαράσταση των πιθανών λύσεων για ένα δεδομένο πρόβλημα σε συνδυασμό με την εφαρμογή σε αυτή μιας οικογένειας γενετικών τελεστών, αποτελεί πολύ χρήσιμο εργαλείο στην προσπάθεια προσέγγισης των πραγματικών

λύσεων σε μια πολύ μεγάλη ποικιλία προβλημάτων και εφαρμογών. Αυτό το γεγονός μετατρέπει αυτή την προσέγγιση "φυσικού μοντέλου" σε μια πολλά υποσχόμενη κατεύθυνση όσον αφορά την επίλυση προβλημάτων γενικότερα.

ΑΣΑΦΗΣ ΛΟΓΙΚΗ

Τον όρο «ασαφή λογική» (fuzzy logic) εισήγαγε το 1965 με άρθρο του ο L.A. Zadeh, (Zadeh, 1965), ο οποίος αναφέρθηκε στην αναγκαιότητα δημιουργίας μίας μαθηματικής θεωρίας που θα επεξεργάζεται ασαφείς-ανακριβείς έννοιες, οι οποίες δεν είναι δυνατό να μοντελοποιηθούν με τη θεωρία των πιθανοτήτων (Zadeh, 1965). Η ανακρίβεια, ή η ασάφεια είναι ο πυρήνας των ασαφών συνόλων και της ασαφούς λογικής. Τα ασαφή σύνολα ουσιαστικά αποτελούν μια γενίκευση των κλασσικών συνόλων. Από τις αρχές της δεκαετίας του 1980 τα ασαφή σύνολα βρήκαν πολλές εφαρμογές, ιδίως σε συστήματα ελέγχου. Ένα ασαφές σύνολο (fuzzy set) A ορίζεται ως ένα σύνολο διατεταγμένων ζευγών $(x, \mu_A(x))$ όπου $x \in X$ και $\mu_A(x) \in [0,1]$. Το σύνολο X αποτελεί ένα ευρύτερο σύνολο αναφοράς (universe of discourse) που περιλαμβάνει όλα τα αντικείμενα στα οποία μπορεί να γίνει αναφορά. Η τιμή $\mu_A(x)$ λέγεται βαθμός αλήθειας, συμβολίζει το βαθμό συγγένειας του x στο A και παίρνει τιμές στο διάστημα $[0,1]$. Τέλος η συνάρτηση μ_A ονομάζεται συνάρτηση συμμετοχής (membership function). Στην πράξη η συνάρτηση συμμετοχής μπορεί να προέρχεται από:

- Υποκειμενικές εκτιμήσεις
- Προκαθορισμένες (ad hoc) και απλοποιημένες μορφές
- Συχνότητες εμφανίσεων και πιθανότητες
- Φυσικές μετρήσεις

- Διαδικασίες μάθησης και προσαρμογής (π.χ. με νευρωνικά δίκτυα)

Η διαφορά των ασαφών συνόλων συγκριτικά με την κλασσική θεωρία συνόλων είναι ότι στην κλασσική θεωρία συνόλων ισχύει $\mu_A(x) \in \{0,1\}$, δηλαδή το x είτε ανήκει στο A ($\mu_A(x)=1$) ή δεν ανήκει ($\mu_A(x)=0$). Άρα η ασαφής θεωρία συνόλων μεταπίπτει στην αντίστοιχη κλασσική, όταν οι δυνατές τιμές της συνάρτησης συμμετοχής είναι 0 ή 1.

Ιδιότητες των Ασαφών Συνόλων

Έστω X ένα σύνολο αντικειμένων, του οποίου τα στοιχεία συμβολίζονται με το γράμμα x . Η συμμετοχή σε ένα υποσύνολο A του συνόλου X είναι μια συνάρτηση συμμετοχής μ_A από το X στο διάστημα $[0,1]$. Το A είναι ένα ασαφές υποσύνολο του X , το οποίο όμως δεν έχει αυστηρά καθορισμένα σύνορα. μ_A είναι ο βαθμός συμμετοχής του στοιχείου x στο A . Όσο πιο κοντά στο 1 είναι το μ_A τόσο πιο πολύ ανήκει το x στο A . Το σύνολο A μπορεί να προσδιοριστεί επομένως από το σύνολο των παρακάτω ζευγών: $A = \{(x, \mu_A(x), x \in X\}$, όπου Κάθε ζευγάρι $(x, \mu_A(x))$ ονομάζεται μονοσύνολο

Χώρος αναφοράς του ασαφούς συνόλου A είναι το σύνολο των στοιχείων του X που έχουν μη μηδενικό βαθμό συμμετοχής στο A .

$$\text{Supp } A = \{x \in X | \mu_A(x) > 0\}.$$

Το σύνολο A μπορεί να γραφεί και ως: $A = \mu_1/x_1 + \mu_2/x_2 + \dots$ ή $A = \sum \mu_i/x_i$

Στο συνεχή χώρο αναφοράς: $A = \int_X \mu_A(x)/x$

Ασαφές δυναμοσύνολο : Ασαφές δυναμοσύνολο , $F(x)$, του υπερσυνόλου αναφοράς X , ονομάζεται το σύνολο όλων των ασαφών υποσυνόλων του X

Υποσύνολο: Το σύνολο A είναι υποσύνολο του B , αν και μόνο αν $\mu_A(x) \leq \mu_B(x)$, για κάθε x ανήκει $X, (x \in X)$

Αν ταυτόχρονα τα A και B δεν είναι ίσα, τότε το A θα ονομάζεται γνήσιο υποσύνολο του B

Ασαφής διαμέριση: Μια οικογένεια ασαφών υποσυνόλων του X , θα λέγεται ασαφής διαμέριση $P^n(X)$ του X τάξης n ($n \in \mathbb{N}$) και θα συμβολίζεται με : $A^n = \{A_1, A_2, \dots, A_n\}$ αν και μόνο αν :

$$A_j \neq A_i, \forall i, j \in N_n (i \neq j)$$

$$0 < \sum_{k=1}^m \mu_{A_i}(x_k) < m, \forall i \in N_n$$

Τα στοιχεία $A_i, i \in N_n$ της A_n θα λέγονται κλάσεις της ασαφούς διαμέρισης.

Κενό ασαφές σύνολο: Ένα ασαφές σύνολο με χώρο αναφοράς τον X , λέγεται κενό αν για κάθε στοιχείο x που ανήκει στον X , η συνάρτηση συμμετοχής του A είναι μηδέν.

$$A = 0 \text{ αν } \mu_A(x) = 0 \forall x \in X$$

Κανονικό ασαφές σύνολο: Ένα ασαφές σύνολο A που ορίζεται στο χώρο αναφοράς X , λέγεται κανονικό αν υπάρχει τουλάχιστο ένα στοιχείο x του X για το οποίο η συνάρτηση συμμετοχής να παίρνει τιμή ίση με τη μονάδα.

$$A = \text{κανονικό αν } \exists x_1: \mu_A(x_1) = 1$$

Ισότητα ασαφών συνόλων: Δύο ασαφή σύνολα A και B που ορίζονται στο χώρο αναφοράς X , λέγονται ίσα αν για κάθε στοιχείο x του X οι συναρτήσεις συμμετοχής των A και B είναι ίσες .

$$\text{Δηλαδή: } A=B \text{ αν } \mu_A(x) = \mu_B(x) \quad \forall x \in X$$

Συστολή ασαφών συνόλων : Έστω ασαφές σύνολο A που ορίζεται στο χώρο αναφοράς X . Η συστολή, $CON(A)$ του συνόλου αυτού είναι ένα νέο ασαφές σύνολο με συνάρτηση συμμετοχής που ορίζεται ως εξής :

$$\mu_{CON(A)}(x) = (\mu_A(x))^2$$

Η συστολή του ασαφούς συνόλου αντιστοιχεί στην προσθήκη του όρου πολύ, μπροστά από τη λεκτική μεταβλητή που περιγράφει το ασαφές σύνολο . Δηλαδή η συστολή του ασαφούς συνόλου ψηλός είναι ένα νέο ασαφές σύνολο που αντιστοιχεί στην έννοια πολύ ψηλός.

Διαστολή ασαφών συνόλων: Αντιστοίχως η διαστολή ενός ασαφούς συνόλου A , είναι ένα νέο ασαφές σύνολο που συμβολίζεται με $DIL(A)$ και έχει συνάρτηση συμμετοχής που περιγράφεται από την παρακάτω σχέση:

$$\mu_{DIL(A)}(x) = \sqrt{\mu_A(x)}$$

Η διαστολή του ασαφούς συνόλου αντιστοιχεί στην προσθήκη του όρου λίγο, μπροστά από τη λεκτική μεταβλητή που περιγράφει το ασαφές σύνολο . Δηλαδή η διαστολή του ασαφούς συνόλου ψηλός είναι ένα νέο ασαφές σύνολο που αντιστοιχεί στην έννοια λίγο ψηλός

Συναρτήσεις συμμετοχής

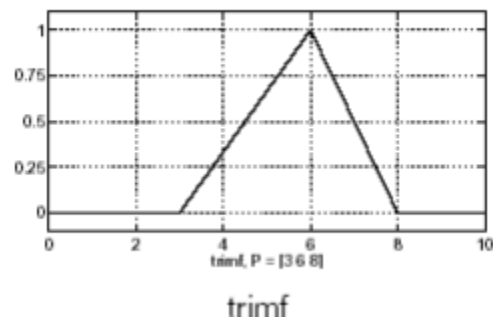
Θα παρουσιάσουμε ορισμένες μορφές μονοδιάστατων παραμετρικών συναρτήσεων συμμετοχής, δηλαδή συναρτήσεων συμμετοχής με μία μοναδική είσοδο (input).

Triangular MFs (συναρτήσεις συμμετοχής τριγωνικής μορφής). Μια τριγωνικής μορφής συνάρτηση συμμετοχής προσδιορίζεται από τρεις παραμέτρους $\{a,b,c\}$ ως ακολούθως :

$$\text{trimf}(x, a, b, c) = \max\left(\min\left(\frac{x-a}{b-a}, \frac{c-x}{c-b}\right), 0\right)$$

Οι παράμετροι $\{a,b,c\}$ (όπου $a < b < c$) καθορίζουν τις συντεταγμένες x των τριών γωνιών της συγκεκριμένης τριγωνικής συνάρτησης συμμετοχής. Η μαθηματική έκφραση της τριγωνικής συνάρτησης συμμετοχής είναι η εξής:

$$\mu = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & x \in (a, b) \\ \frac{c-x}{c-b}, & x \in (b, c) \\ 0, & x \geq c \end{cases}$$



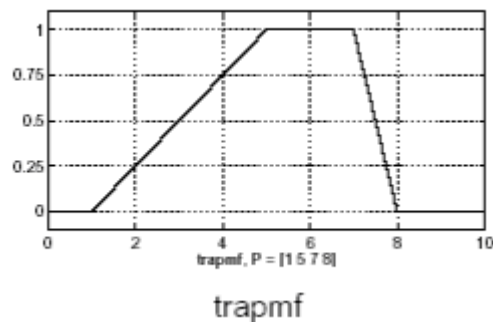
Trapezoidal MFs:

Η τραπεζοειδής συνάρτηση συμμετοχής προσδιορίζεται από τέσσερις παραμέτρους $\{a,b,c,d\}$ ως ακολούθως:

$$\mu_{\text{trimf}}(x, a, b, c, d) = \max \left(\min \left(\frac{x-a}{b-a}, 1, \frac{d-x}{d-c} \right), 0 \right)$$

Οι παράμετροι $\{a,b,c,d\}$ με $a < b < c < d$ καθορίζουν τις συντεταγμένες x των τεσσάρων γωνιών της συγκεκριμένης τραπεζοειδούς συνάρτησης συμμετοχής. Παρακάτω δίδεται η μαθηματική έκφραση της τραπεζοειδούς συνάρτησης συμμετοχής:

$$A = \begin{cases} 0, x \leq a \\ \frac{x-a}{b-a}, x \in (a, b) \\ 1, x \in (b, c) \\ \frac{d-x}{d-c}, x \in (c, d) \\ 0, x \geq d \end{cases}$$

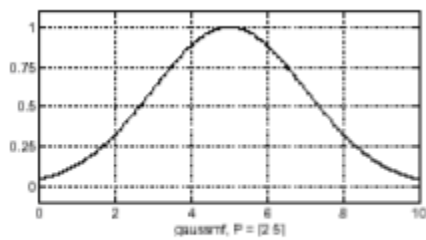


Gaussian MFs:

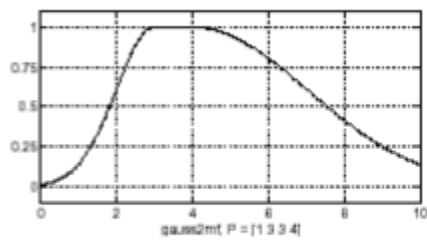
Μια Γκαουσιανή συνάρτηση συμμετοχής προσδιορίζεται από δύο παραμέτρους

$$gaussmf(x, c, \sigma) = e^{-\frac{1}{2} \left(\frac{x-c}{\sigma} \right)^2}$$

$\{c, \sigma\}$



gaussmf



gauss2mf

Η Γκαουσιανή συνάρτηση συμμετοχής προσδιορίζεται επακριβώς από δύο παραμέτρους $\{c, \sigma\}$.

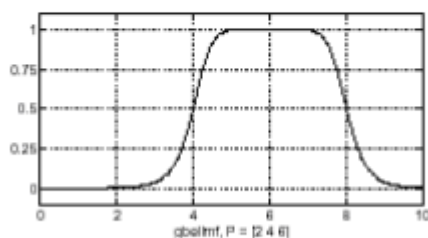
Το c αναπαριστά το κέντρο της, και το σ το πλάτος της.

Η Γκαουσιανή και η καμπάνα μπορούν να χρησιμοποιούνται συχνά στα ασαφή σύνολα λόγω της ομαλότητάς τους. Έχουν δε το πλεονέκτημα να διατηρούν μη μηδενικές τιμές σε όλα τα σημεία.

Generalized bell MFs:

Η γενικευμένη καμπανοειδής συνάρτηση συμμετοχής προσδιορίζεται από τρεις παραμέτρους $\{a, b, c\}$, όπου η παράμετρος b είναι συνήθως θετική.

$$gbellmf(x, a, b, c) = \frac{1}{1 + \left| \frac{x - c}{a} \right|^{2b}}$$

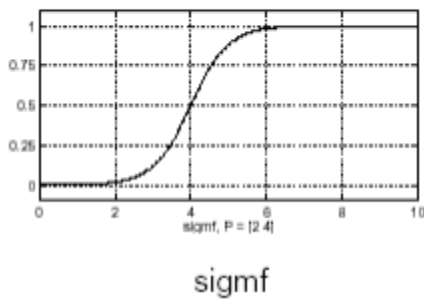


gbellmf

Sigmoidal MF

Η σιγμοειδής συνάρτηση συμμετοχής προσδιορίζεται ως ακολούθως:

$$\text{sigmf}(x, a, c) = \frac{1}{1 + \exp[-a(x - c)]}$$



Το a ρυθμίζει την κλίση στο crossover point $x=c$. Ανάλογα με την τιμή που θα πάρει το a η σιγμοειδής MF είναι open left, open right και επομένως είναι κατάλληλη να χρησιμοποιηθεί στην αναπαράσταση εννοιών όπως «πολύ μεγάλο» ή «πολύ αρνητικό». Σιγμοειδείς συναρτήσεις αυτής της μορφής χρησιμοποιούνται ευρύτατα ως συναρτήσεις ενεργοποίησης (activation function) των τεχνητών νευρωνικών δικτύων.

Λογικές Πράξεις στα Ασαφή Σύνολα

Μέχρι τώρα έχουμε αναφερθεί στα ασαφή σύνολα και καθόλου στην ασαφή λογική. Η ασαφής λογική δεν είναι παρά ένα υπερσύνολο της λογικής Boolean. Αν δηλαδή απομονώσουμε τους ακραίους βαθμούς συμμετοχής 0 (πλήρως ψευδές) και 1 (πλήρως αληθές) τότε οι κλασσικοί λογικοί τελεστές μπορούν να εφαρμοσθούν. Για παράδειγμα ας θεωρήσουμε τον παρακάτω πίνακα αληθείας:

| A | B | A and B |
|---|---|---------|
| 0 | 0 | 0 |
| 0 | 1 | 0 |
| 1 | 0 | 0 |
| 1 | 1 | 1 |

AND

| A | B | A or B |
|---|---|--------|
| 0 | 0 | 0 |
| 0 | 1 | 1 |
| 1 | 0 | 1 |
| 1 | 1 | 1 |

OR

| A | not A |
|---|-------|
| 0 | 1 |
| 1 | 0 |

NOT

Σχήμα : Πίνακας αληθείας Boolean

Μεταβαίνοντας στον χώρο της ασαφούς λογικής πρέπει να έχουμε υπόψη, ότι οι έννοιες αληθές και ψευδές είναι θέμα βαθμού συμμετοχής. Επομένως ο πίνακας αυτός πρέπει να μετατραπεί με τρόπο που να συμπεριλαμβάνει αυτήν την αρχή. Οι τιμές των εισόδων A και B είναι τώρα πραγματικοί αριθμοί από το 0 μέχρι το 1. Πρέπει λοιπόν να ευρεθεί μια συνάρτηση που να διατηρεί τις ιδιότητες της συνάρτησης AND και ταυτόχρονα να μπορεί να επεκτείνεται για πραγματικούς αριθμούς.

Μια πιθανή απάντηση μπορεί να είναι ο τελεστής $\min(A,B)$, η ελάχιστη δηλαδή τιμή των εισόδων A και B. Με βάση το ίδιο σκεπτικό μια συνάρτηση που μπορεί να αντικαταστήσει τον τελεστή OR της Boolean λογικής είναι ο τελεστής $\max(A,B)$. Τέλος ο τελεστής NOT A μπορεί να αντικατασταθεί με την πράξη $1-A$.

Παρακάτω βλέπουμε ότι ο πίνακας αληθείας παραμένει αμετάβλητος αν εφαρμόσουμε τις παραπάνω υποκαταστάσεις.

| A | B | $\min(A,B)$ |
|---|---|-------------|
| 0 | 0 | 0 |
| 0 | 1 | 0 |
| 1 | 0 | 0 |
| 1 | 1 | 1 |

AND

| A | B | $\max(A,B)$ |
|---|---|-------------|
| 0 | 0 | 0 |
| 0 | 1 | 1 |
| 1 | 0 | 1 |
| 1 | 1 | 1 |

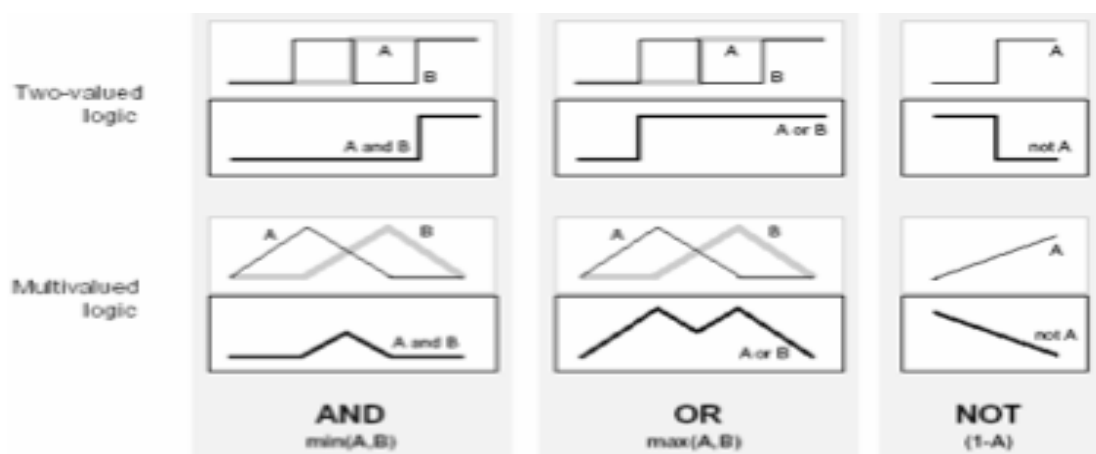
OR

| A | $1-A$ |
|---|-------|
| 0 | 1 |
| 1 | 0 |

NOT

Σχήμα : Πίνακας αληθείας ασαφούς λογικής

Αφού έχουμε ορίσει συναρτήσεις που μπορούν ορίσουν τον πίνακα αληθείας, μπορούμε πλέον να επεκταθούμε και στην περίπτωση πραγματικών αριθμών. Στο παρακάτω σχήμα έχουμε αντικαταστήσει τον πίνακα αληθείας με ένα γράφημα δύο ασαφών συνόλων. Στο πάνω μέρος έχουμε την περίπτωση συνόλων με δύο τιμές ενώ στο κάτω φαίνεται ο τρόπος που φαίνεται πως λειτουργούν οι τελεστές στην περίπτωση που οι τιμές αληθείας A και B μεταβάλλονται συνεχώς από το 0 στο 1.



Σχήμα : Κατασκευή νέων καμπυλών από ήδη υπάρχουσες, με την βοήθεια τελεστών.

Δεδομένων των παραπάνω συναρτήσεων μπορούμε να κατασκευάσουμε δομές με βάση ασαφή σύνολα και τους λογικούς κανόνες AND OR και NOT . Βέβαια το γεγονός ότι βρήκαμε συναρτήσεις που να επεκτείνουν τους τελεστές αυτούς από το χώρο της Boolean λογικής στο χώρο της ασαφούς λογικής δε σημαίνει σε καμία περίπτωση ότι οι συναρτήσεις αυτές είναι και οι μοναδικές.

Έχουμε ορίσει την τομή , την ένωση και το συμπλήρωμα ενός ασαφούς συνόλου με τους τελεστές \min \max $1-A$ οι οποίοι είναι οι πιο συχνά χρησιμοποιούμενοι.

Ασαφείς Σχέσεις

Οι ασαφείς σχέσεις (fuzzy relations) είναι ασαφή σύνολα ορισμένα σε πεδία αναφοράς ανώτερης διάστασης (π.χ. $X \times X$, $X \times Y \times Z$ κλπ). Ποιοτικά , μια ασαφής σχέση R θα μπορούσε να είναι μια έκφραση της μορφής «είναι βαρύτερο από » και η οποία θα συνδέει τα στοιχεία δύο ή άλλων συνόλων:

$R = \{ \langle x, y \rangle \mid x \in X, y \in Y \text{ και } \mu_R(x, y) > 0 \}$

Οι ασαφείς σχέσεις μπορεί να εκφραστούν με αναφορά όλων των ζευγών (τιμή, βαθμός συμμετοχής) , δηλαδή ζευγών της μορφής $((x, y), \mu_R(x, y))$. Ένας άλλος τρόπος αναπαράστασης , ιδιαίτερα χρήσιμος σε υπολογισμούς είναι σε μορφή

$$R = \begin{bmatrix} \mu_R(x_1, y_1) & \mu_R(x_1, y_2) & \cdots & \mu_R(x_1, y_n) \\ \mu_R(x_2, y_1) & \mu_R(x_2, y_2) & \cdots & \mu_R(x_2, y_n) \\ \vdots & \vdots & \ddots & \vdots \\ \mu_R(x_m, y_1) & \mu_R(x_m, y_2) & \cdots & \mu_R(x_m, y_n) \end{bmatrix}$$

πίνακα:

Οι ασαφείς σχέσεις μπορούν να συνδυαστούν μεταξύ τους μέσω της διαδικασίας της σύνθεσης (composition). Αν για παράδειγμα συνδυαστεί η ασαφής σχέση $R_1(x,y)$ ορισμένη στο $X \times Y$ με την ασαφή σχέση $R_2(y,z)$ ορισμένη στο $Y \times Z$ τότε θα προκύψει μία ασαφής σχέση $R(x,z)$ η οποία θα ορίζεται στο σύνολο $X \times Z$ και θα συσχετίζει άμεσα στοιχεία των συνόλων X και Z . Βέβαια είναι απαραίτητο να προσδιοριστεί επακριβώς η συνάρτηση συμμετοχής $\mu_{R(x,z)}$ της R με χρήση των συναρτήσεων συμμετοχής των R_1 και R_2 .

Οι βασικές πράξεις που ορίζονται μεταξύ των ασαφών σχέσεων είναι η αντιστροφή και η σύνθεση.

Αντίστροφη σχέση της $R(X,Y)$ είναι η ασαφής σχέση $R^{-1}(X,Y)$ με τύπο :

$R^{-1}(x,y)=R(y,x)$ για κάθε x που ανήκει στον X και για κάθε y που ανήκει στον Y .

Η σύνθεση είναι πολύ σημαντική διαδικασία. Οι περισσότεροι γνωστές μέθοδοι σύνθεσης ασαφών συνόλων είναι η σύνθεση max-min composition και η σύνθεση max-product composition. Αν $R_1(x,y)$ και $R_2(y,z)$ είναι δύο ασαφείς σχέσεις ορισμένες στα σύνολα $X \times Y$ και $Y \times Z$ αντίστοιχα, τότε η σύνθεση τους δίνει μια σχέση $R_1 \circ R_2$ ορισμένη στο $X \times Z$.

Ασαφής Συλλογιστική

Η βάση στην οποία στηρίζεται η λήψη αποφάσεων είναι η παραγωγή συλλογιστικής. Η ασαφής λογική ασχολείται με την παραγωγή συλλογιστικής σε περιβάλλον αβεβαιότητας. Για το σκοπό αυτό, θεμελιώνεται η δομή και η μαθηματική αναπαράσταση ενός ασαφούς γεγονότος με τον ορισμό των ασαφών συνόλων και καθορίζεται ο τρόπος με τον οποίο συνδυάζουμε τα γεγονότα για να παράγουμε λογικές προτάσεις ή σχέσεις και συνεπώς συμπεράσματα.

Οι συλλογιστικοί τρόποι που κυρίως χρησιμοποιούνται είναι τρεις:

- Ο modus ponens (MP)
- Ο modus tolens (MT)
- υποθετικός συλλογισμός (HS)

Ο modus ponens παράγει συμπεράσματα από ένα σύνολο υποθέσεων σύμφωνα με το σχήμα:

$$(A \Rightarrow B) \wedge A \Rightarrow B$$

όπου A και B συγκεκριμένα γεγονότα.

Ο παραπάνω κανόνας ερμηνεύεται ως εξής : Αν το γεγονός A συνεπάγεται το γεγονός B και επιπλέον έχουμε ως υπόθεση ότι ισχύει το A , τότε το συμπέρασμα που παίρνουμε είναι ότι ισχύει και το B . Όμως σε περιβάλλον ασάφειας τα γεγονότα ισχύουν σε κάποιο βαθμό . Έτσι ο παραπάνω κανόνας πρέπει να τροποποιηθεί για να συμπεριλάβει και την ασάφεια . Καταλήγουμε με αυτό τον τρόπο στο γενικευμένο κανόνα modus ponens (generalized modus ponens) ο οποίος έχει την παρακάτω μορφή:

$$(A \Rightarrow B) \wedge A' \Rightarrow B'$$

Αντίστοιχα ο γενικευμένος κανόνας modus tolens (generalized modus tolens) διατυπώνεται ως εξής:

$$(A \Rightarrow B) \wedge B' \Rightarrow A'$$

Ας εξετάσουμε τώρα την ερμηνεία των παραπάνω κανόνων . Για παράδειγμα ο γενικευμένος κανόνας modus ponens ερμηνεύεται ως εξής : Αν το γεγονός A συνεπάγεται το γεγονός B και έχουμε ως υπόθεση ότι ισχύει το A σε κάποιο βαθμό, τότε θα ισχύει και το B σε κάποιο βαθμό.

Πρέπει, λοιπόν να προσδιοριστεί ο βαθμός στον οποίο πληρείται το γεγονός B . Αυτός εξαρτάται από το βαθμό στον οποίο πληρείται το γεγονός A και από το είδος της συνεπαγωγής που εφαρμόζουμε . Η πράξη της ασαφούς συνεπαγωγής υλοποιεί τη μαθηματική σχέση $A \Rightarrow B$, όταν τα A και B είναι ασαφή γεγονότα..

Το σχήμα που προτείνεται από τη θεωρία των ασαφών συνόλων για την εξαγωγή του γεγονότος B' από τα γεγονότα A , A', B περιγράφεται από την παρακάτω σχέση που αποτελεί το συνθετικό κανόνα του Zadeh:

$$B'(y) = \sup_{x \in X} [A(x), \sigma(A(x), B(y))]$$

όπου σ είναι μια συνάρτηση που υλοποιεί την πράξη της ασαφούς συνεπαγωγής . Επομένως η επιλογή της συνάρτησης που υλοποιεί την ασαφή συνεπαγωγή έχει ουσιαστικό ρόλο για την ασαφή συλλογιστική που παράγεται με βάση το παραπάνω σχήμα. Ένα από τα κριτήρια που χρησιμοποιείται για αυτή την επιλογή είναι το κριτήριο της ανάκλησης (recall) , το οποίο διατυπώνεται ως εξής:

$$B(y) = \sup_{x \in X} [A(x), \sigma(A(x), B(y))]$$

ο κριτήριο αυτό έχει την εξής ερμηνεία: η ασαφής συνεπαγωγή πρέπει να είναι τέτοια ώστε όταν η υπόθεση πληρείται ακριβώς , τότε να λαμβάνουμε το συμπέρασμα του κανόνα $A \Rightarrow B$, δηλαδή το γεγονός B . η απαίτηση αυτή είναι εύλογη αφού όταν δεν υπάρχει αβεβαιότητα η ασαφής συλλογιστική οφείλει να ταυτίζεται με την κλασσική συλλογιστική.

Συστήματα Ασαφούς Συλλογιστικής-Νευροασαφή συστήματα

Εισαγωγή

Στο προηγούμενο κεφάλαιο αναφερθήκαμε στα ασαφή σύνολα και στις πράξεις που γίνονται σ' αυτά. Στο παρόν κεφάλαιο θα αναφέρουμε τον τρόπο που μπορούν να εφαρμοσθούν τα προηγούμενα σε ένα σύστημα λήψης αποφάσεων. Θα εξετάσουμε ένα απλοϊκό παράδειγμα λήψης αποφάσεων και θα δούμε πως μπορεί να λυθεί με τη βοήθεια της ασαφούς λογικής. Εξετάζουμε το πρόβλημα του φιλοδώρηματος, πόσο δηλαδή είναι το φιλοδώρημα που πρέπει να δώσει ένας πελάτης σε ένα εστιατόριο.

Οι παράμετροι του προβλήματος είναι η ποιότητα του φαγητού και η εξυπηρέτηση.

Είναι προφανές για παράδειγμα ότι αν το φαγητό είναι άριστης ποιότητας και η εξυπηρέτηση καλή τότε το φιλοδώρημα θα είναι υψηλό. Αντίστοιχα αν η ποιότητα της εξυπηρέτησης είναι χαμηλή και το φαγητό κακό τότε το φιλοδώρημα θα είναι χαμηλό.

Τι γίνεται όμως αν έχουμε μια ενδιάμεση κατάσταση, π.χ. μέτριο φαγητό και καλή εξυπηρέτηση.

Επιπλέον παραπάνω έχουμε αναφέρει προσδιορισμούς όπως καλό, καλό, μέτριο που εμπεριέχουν κάποια ασάφεια.

Αν προσπαθήσουμε να λύσουμε το πρόβλημα γραμμικά θα πρέπει να δώσουμε καθορισμένους κανόνες για κάθε μια περίπτωση. Μέσω όμως της ασαφούς λογικής το πρόβλημα αυτό μπορεί να λυθεί ευκολότερα. Πριν προχωρήσουμε στη λύση του προβλήματος θα δούμε πως μπορούμε να υλοποιήσουμε τους κανόνες αν-τότε (if-then rules) που προαναφέραμε (αν το φαγητό είναι κακό κτλ).

Κανόνες αν-τότε (if-then rules)

Τα ασαφή σύνολα και οι ασαφείς τελεστές είναι τα υποκείμενα και τα ρήματα του ασαφούς λογισμού. Οι προτάσεις if-then χρησιμοποιούνται για να σχηματίσουν τις συνθήκες εκείνες που συνιστούν την ασαφή λογική. Ένας απλός κανόνας if-then έχει τη μορφή:

if x is A then y is B

, όπου A και B είναι οι γλωσσικές μεταβλητές που προσδιορίζονται από ασαφή σύνολα με χώρο αναφοράς X και Y αντίστοιχα. Το πρώτο κομμάτι του κανόνα 'if x is A' ονομάζεται υπόθεση (antecedent ή premise) ενώ το δεύτερο κομμάτι 'then y is B' ονομάζεται συνέπεια ή συμπέρασμα (consequent ή conclusion). Ένα παράδειγμα ενός τέτοιου κανόνα μπορεί να είναι το εξής:

if service is good then tip is average

Η λέξη good αντιπροσωπεύεται από έναν αριθμό που κυμαίνεται ανάμεσα στο 0 και το 1. Έτσι η υπόθεση είναι μια μετάφραση που επιστρέφει μια τιμή από το 0 έως το 1. Από την άλλη μεριά η λέξη average αποτελεί ένα ασαφές σύνολο και έτσι η συνέπεια του κανόνα είναι μια συσχέτιση της εξόδου y στο ασαφές σύνολο B.

Στους κανόνες if-then η λέξη 'is' έχει διαφορετική έννοια ανάλογα με το αν εμφανίζεται στο πρώτο ή στο δεύτερο μέρος του κανόνα. Έτσι όταν εμφανίζεται στο πρώτο μέρος του κανόνα έχει την έννοια του ελέγχου ισότητας (δηλαδή την ίδια έννοια που δίνουμε στο σύμβολο $=$).

Αντίθετα όταν εμφανίζεται στο δεύτερο μέρος του κανόνα έχει την έννοια της καταχώρησης δηλαδή την έννοια που δίνουμε στο σύμβολο $=$). Ο παραπάνω κανόνας λοιπόν θα μπορούσε να γραφεί συμβολικά και ως εξής:

if service == good then tip = average

Γενικά η είσοδος σε έναν κανόνα if-then είναι η τρέχουσα τιμή της μεταβλητής εισόδου, ενώ η έξοδος του κανόνα είναι ολόκληρο το ασαφές σύνολο. Από αυτό όμως το ασαφές σύνολο θα πρέπει να κρατήσουμε μια τιμή, έτσι ώστε να μπορούμε να βγάλουμε κάποια συμπεράσματα. Γι'

αυτό πρέπει να γίνει μια διαδικασία που ονομάζεται αποσαφήνιση (defuzzification). Για την αποσαφήνιση αναφερόμαστε εκτενέστερα παρακάτω. Για να διερμηνεύσουμε έναν κανόνα If-then πρέπει να ακολουθήσουμε κάποια στάδια:

1. Πρώτα πρέπει να υπολογίσουμε την είσοδο. Για να γίνει αυτό όμως πρέπει να γίνει ασαφοποίηση της εισόδου και να εφαρμοσθούν οι κατάλληλοι ασαφείς τελεστές

2. Να εφαρμόσουμε το παραπάνω αποτέλεσμα στο δεύτερο μέρος του κανόνα

Αν πάρουμε την απλή περίπτωση των δυαδικών τιμών οι κανόνες if-then είναι αρκετά εύκολο να εφαρμοσθούν. Αν η προϋπόθεση είναι αληθής τότε και η συνέπεια θα είναι αληθής. Πώς όμως επεκτείνεται η παραπάνω συλλογιστική για να συμπεριλάβει και ασαφείς τιμές; Όταν η προϋπόθεση είναι αληθής σε κάποιο βαθμό τότε και η συνέπεια είναι αληθής στον ίδιο βαθμό.

Με άλλα λόγια:

Στη δυαδική λογική: $p \rightarrow q$ (Τα p και q είναι και τα δύο είτε πλήρως αληθή ή πλήρως ψευδή)

Στην ασαφή λογική: $0.5 p \rightarrow 0.5 q$ (Τα p και q είναι μερικώς αληθή και μερικώς ψευδή)

Η προϋπόθεση ενός κανόνα μπορεί να αποτελείται από περισσότερα του ενός μέρη. Όπως για παράδειγμα ο παρακάτω κανόνας:

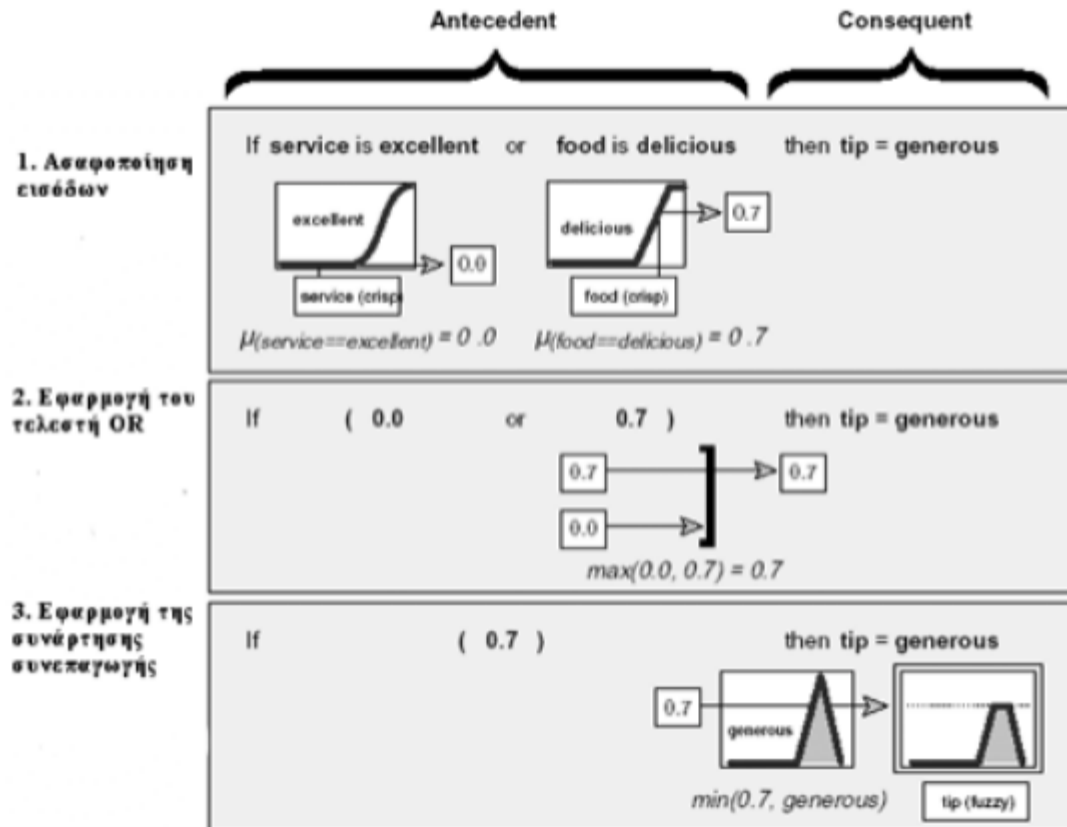
if sky is gray and wind is strong and barometer is falling, then .

Σ' αυτή την περίπτωση όλα τα μέρη της προϋπόθεσης υπολογίζονται ταυτόχρονα και καταλήγουμε σε ένα μοναδικό νούμερο ανάλογο με τους λογικούς τελεστές που χρησιμοποιούμε.

Αλλά και η συνέπεια ενός κανόνα μπορεί να αποτελείται από περισσότερα του ενός μέρη . Για παράδειγμα ο παρακάτω κανόνας:

if temperature is cold then hot water valve is open and cold water valve is shut

Όλα τα μέρη της συνέπειας του κανόνα επηρεάζονται το ίδιο από το αποτέλεσμα της προϋπόθεσης. Η συνέπεια του κανόνα προσδιορίζει ένα ασαφές σύνολο το οποίο συσχετίζεται με την έξοδο . Η συνάρτηση συνεπαγωγής (implication function) στη συνέχεια τροποποιεί το ασαφές σύνολο στο βαθμό που προσδιορίζεται από την προϋπόθεση του κανόνα .Μια πολύ συνηθισμένη συνάρτηση για να γίνει αυτό είναι η αποκοπή με τη χρήση του τελεστή min. Στο παρακάτω σχήμα βλέπουμε τον τρόπο με τον οποίο εφαρμόζονται οι κανόνες if-then.



Συστήματα ασαφούς συλλογιστικής (fuzzy inference systems)

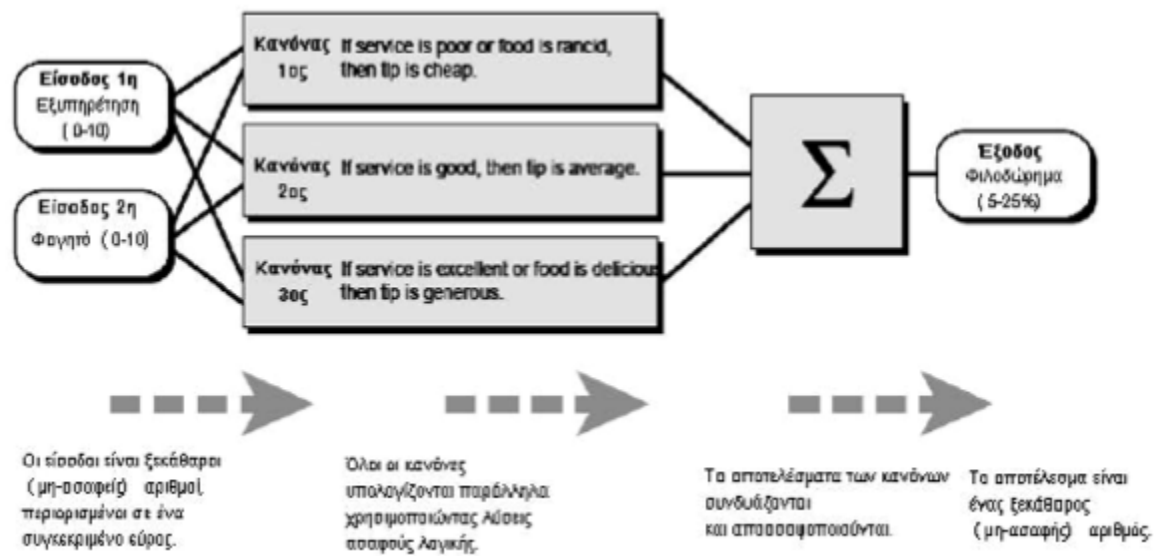
Γενικά ένας κανόνας από μόνος του δεν επαρκεί για τα περισσότερα πρακτικά προβλήματα. Αυτό που χρειαζόμαστε είναι περισσότεροι του ενός κανόνες οι οποίοι να αλληλεπιδρούν μεταξύ τους. Η έξοδος κάθε κανόνα θα είναι ένα ασαφές σύνολο. Τα ασαφή σύνολα που προκύπτουν σε συνέπεια των κανόνων συγκεντρώνονται για να σχηματίσουν ένα ασαφές σύνολο εξόδου από το οποίο μέσω της διαδικασίας της αποσαφήνισης θα πάρουμε μια μοναδική τιμή. Παρακάτω αναπτύσσεται ο τρόπος με τον οποίο εξελίσσεται η παραπάνω διαδικασία μέσα από ένα σύστημα ασαφούς συλλογιστικής (fuzzy inference system, FIS) που ονομάζεται Mamdani.

Συστήματα τύπου Mamdani

Η ασαφής συλλογιστική είναι η διαδικασία κατά την οποία σχηματίζεται η απεικόνιση από μια δεδομένη είσοδο σε μια έξοδο χρησιμοποιώντας τις αρχές της ασαφούς λογικής. Η απεικόνιση αυτή θέτει τις βάσεις με τις οποίες μπορεί να ληφθεί μια απόφαση ή να διακριθούν πρότυπα (patterns). Η διαδικασία αυτή της ασαφούς συλλογιστικής περιλαμβάνει όλα εκείνα τα κομμάτια στα οποία αναφερθήκαμε προηγουμένως. Δηλαδή τις συναρτήσεις συμμετοχής, τους ασαφείς λογικούς τελεστές και τους κανόνες if-then. Υπάρχουν δύο είδη ασαφών συστημάτων: των παρεμβολής τα τύπου -Mamdani και τα τύπου Sugeno. Οι διαφορές ανάμεσα στα δύο αυτά συστήματα έγκεινται κυρίως στον τρόπο με τον οποίο σχηματίζεται η έξοδος του συστήματος. Τα συστήματα Mamdani είναι αυτά που συναντώνται συχνότερα. Η μεθοδολογία των συστημάτων Mamdani εισήχθη το 1975 από τον Ebrahim Mamdani.

Θα εξετάσουμε τώρα πως λειτουργεί ένα FIS τύπου Mamdani μέσα από το πρόβλημα του φιλοδωρήματος που αναφέραμε παραπάνω.

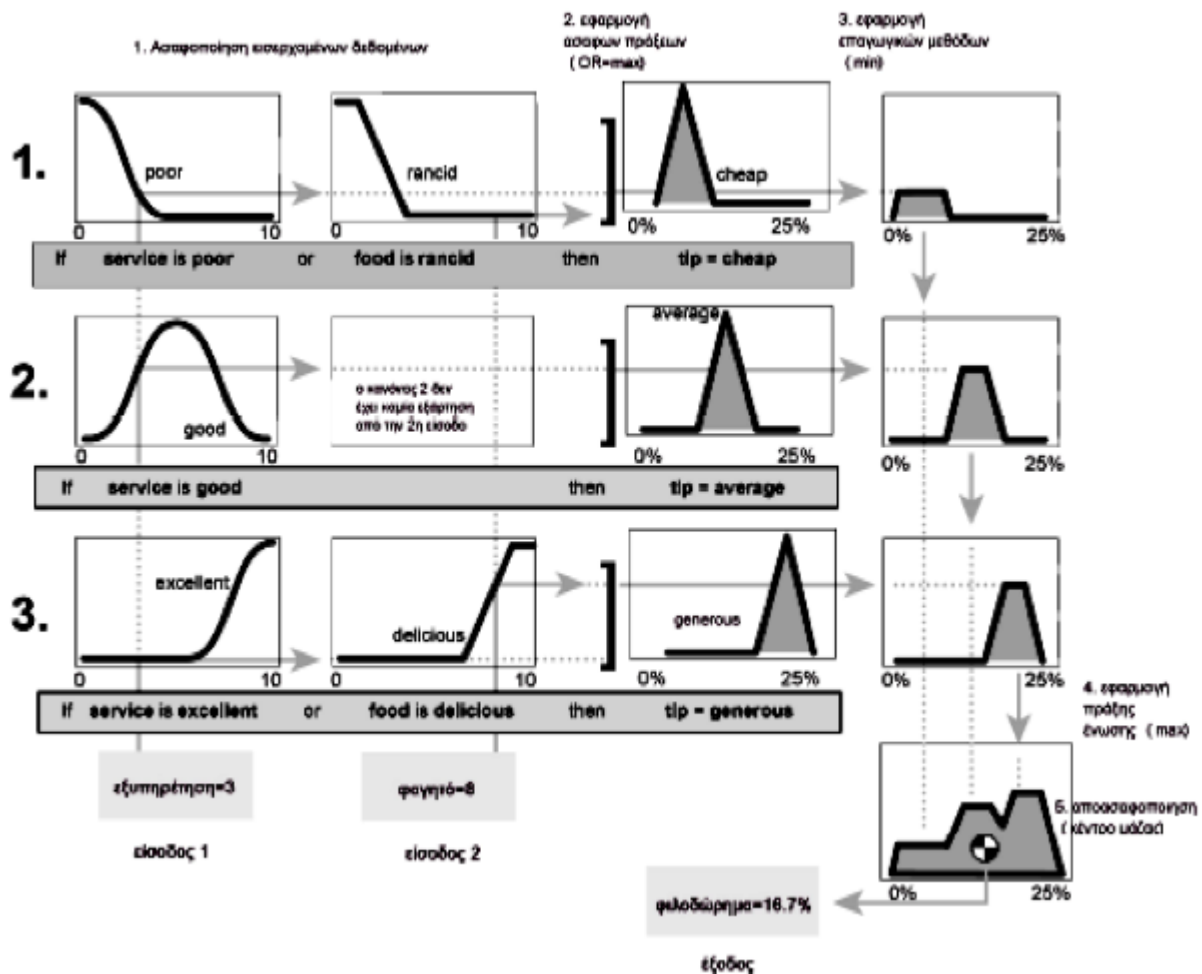
Το σύστημα έχει δύο εισόδους και τρεις κανόνες. Οι είσοδοι του συστήματος είναι η ποιότητα του φαγητού και η εξυπηρέτηση. Θεωρούμε ότι ο πελάτης βαθμολογεί την ποιότητα του φαγητού και την εξυπηρέτηση με μια κλίμακα από το 0 έως το 10. η έξοδος του συστήματος είναι το φιλοδώρημα που θα δώσει ο πελάτης και το οποίο κυμαίνεται από 5% έως 25% της τιμής του λογαριασμού.



Σχήμα : σύστημα 2 εισόδων , 3 κανόνων , 1 εξόδου .

Η πληροφορία οδεύει από τα αριστερά προς τα δεξιά και καταλήγει σε μία μοναδική έξοδο . Ένα σημαντικό χαρακτηριστικό των συστημάτων ασσφούς λογικής είναι η παράλληλη εφαρμογή των κανόνων. Στα κλασσικά συστήματα δυαδικής λογικής έχουμε χρήση διακοπών που ανοίγουν και κλείνουν ανάλογα με τις τιμές των εισόδων . Στα ασσφή συστ ήματα μεταβαίνουμε ομαλά από περιοχή σε περιοχή ανάλογα με τις συνέπειες των κανόνων . Η διαδικασία της απεικόνισης των εισόδων στις εξόδου μέσα από ένα σύστημα ασσφούς λογικής ακολουθεί πέντε βήματα:

1. Ασασφοποίηση των εισόδων.
2. Εφαρμογή των ασσφών λογικών τελεστών στις προϋπόθεσης των κανόνων.
3. Καθορισμός των συνεπαγωγών των κανόνων.
4. Συγκέντρωση όλων των συνεπαγωγών όλων των κανόνων.
5. Αποσασφήνιση.



Σχήμα : Fis Mamdani για ποσοστό φιλοδωρήματος

Συστήματα τύπου Sugeno

Παραπάνω ασχοληθήκαμε με τα συστήματα Mamdani που είναι και τα πιο ευρέως χρησιμοποιούμενα. Ωστόσο υπάρχει και η μέθοδος Sugeno που εισήχθηκε το 1985 η μέθοδος αυτή έχει ομοιότητες με τη μέθοδο των συστημάτων Mamdani. Για παράδειγμα τα 2 πρώτα βήματά της (ασαφοποίηση των εισόδων και εφαρμογή των τελεστών) είναι ακριβώς τα ίδια. Η κύρια διαφορά ανάμεσα στα δύο συστήματα έγκειται στο ότι οι συναρτήσεις συμμετοχής στην έξοδο των συστημάτων Sugeno είναι μόνο γραμμική ή σταθερά.

Ένας τυπικός ασαφής κανόνας σε συστήματα sugeno μηδενικής τάξης έχει την μορφή:

if x is A and y is B then $z = k$,

όπου A και B είναι τα ασαφή σύνολα της προϋπόθεσης ενώ k είναι μια αριθμητική τιμή. Αφού η συνέπεια του κανόνα είναι μια σταθερά τότε αυτό σημαίνει ότι το βήμα 3 εκφυλίζεται σε ένα απλό πολλαπλασιασμό ενώ το βήμα 4 καταλήγει να είναι η συγκέντρωση όλων των σταθερών.

Ένα σύστημα sugeno πρώτης τάξης θα έχει κανόνες με τη γενική μορφή

if x is A and y is B then $z = p \cdot x + q \cdot y + r$,

όπου A και B είναι τα ασαφή σύνολα της προϋπόθεσης ενώ τα p,q,r είναι σταθερές.

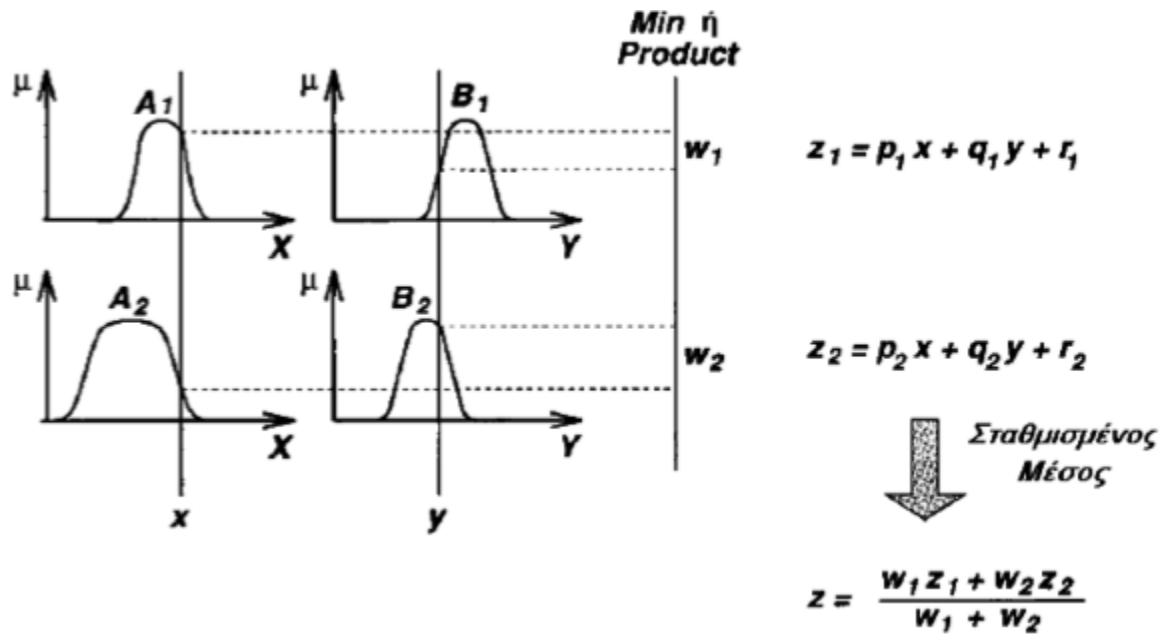
Ένας τρόπος για να δούμε τα συστήματα πρώτης τάξης είναι να θεωρήσουμε ότι κάθε κανόνας προσδιορίζει τη θέση ενός κινούμενου singleton. Το singleton αυτό μπορεί να κινείται στο χώρο της εξόδου, με γραμμικό τρόπο και η θέση του εξαρτάται από τις τιμές των εισόδων.

Συστήματα sugeno ανώτερης τάξης είναι εφικτά, όμως δεν προσφέρουν σημαντικές βελτιώσεις και ταυτόχρονα εισαγάγουν σημαντική πολυπλοκότητα.

Για ένα Sugeno μηδενικού βαθμού το επίπεδο εξόδου z είναι μια σταθερά ($a=b=0$) και μπορεί να θεωρηθεί σαν μια ειδική περίπτωση, στο οποίο το συμπέρασμα του κάθε κανόνα προσδιορίζεται

108

από ένα ασαφές προαποσαφοποιημένο συμπέρασμα (singleton). Η έξοδος ενός Sugeno μηδενικού βαθμού είναι μια ομαλή συνάρτηση των μεταβλητών εισόδων του, εφόσον οι γειτονικές συναρτήσεις συμμετοχής στις προϋποθέσεις έχουν αρκετή επικάλυψη.



Σχήμα : Ασαφές μοντέλο Sugeno πρώτου βαθμού

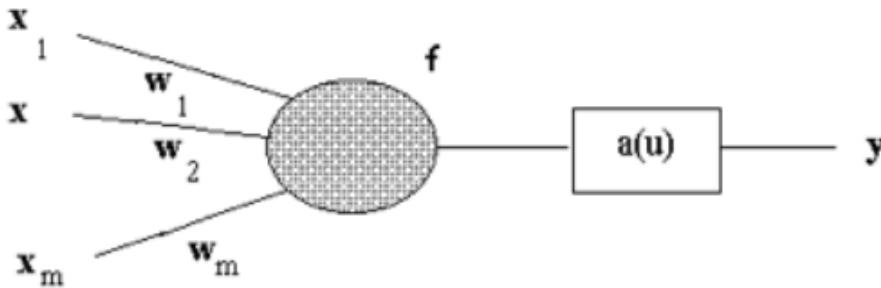
Στο προηγούμενο σχήμα βλέπουμε ότι το επίπεδο εξόδου Z_i για κάθε κανόνα είναι σταθμισμένο με βάρη W_i για κάθε κανόνα.

Σύγκριση Sugeno-Mamdani

| Πλεονεκτήματα Sugeno | Πλεονεκτήματα Mamdani |
|---|--|
| Επαρκέστερα υπολογιστικά | Είναι διαισθητικά |
| Καλή λειτουργία για γραμμικές τεχνικές | Έχουν μεγάλη αποδοχή |
| Καλή λειτουργία για βελτιστοποίηση και προσαρμοστικές τεχνικές | Ευπροσάρμοστα στον ανθρώπινο τρόπο εισαγωγής δεδομένων εισόδου |
| Καλά προσαρμοσμένο στη μαθηματική ανάλυση | |
| Εγγυημένη συνέχεια της τελικής επιφάνειας που αποτελεί την λύση | |

Νευροασαφή συστήματα

ΑΣΑΦΕΙΣ ΝΕΥΡΩΝΕΣ



Όπου:

x_1, x_2, \dots, x_m οι είσοδοι του νευρώνα

w_1, w_2, \dots, w_m τα βάρη των συνάψεων

f η συνάρτηση συμμετοχής του νευρώνα

a η συνάρτηση ενεργοποίηση του νευρώνα

y η έξοδος του νευρώνα

Στην ενότητα αυτή εξετάζουμε τα νευροασαφή συστήματα . Η δομική μονάδα των συστημάτων αυτών είναι οι ασαφείς νευρώνες . Οι ασαφείς νευρώνες διακρίνονται σε νευρώνες συμμετοχής και σε λειτουργικούς νευρώνες . Παρακάτω γίνεται αναφορά στα χαρακτηριστικά κάθε κατηγορίας ασαφών νευρώνων.

ΝΕΥΡΩΝΕΣ ΣΥΜΜΕΤΟΧΗΣ

Οι συνηθέστερες μορφές συναρτήσεων συμμετοχής είναι η τριγωνική, η γκαουσιανή, η τραπεζοειδής , η κανονική κ .ο.κ.. Οι παραπάνω συναρτήσεις συμμετοχής μπορούν να υλοποιηθούν με τη χρήση ενός νευρώνα . Η συνάρτηση συμμετοχής μπορεί να υλοποιηθεί

θεωρώντας ότι ο νευρώνας δέχεται ως είσοδο το x και δίνει έξοδο $A(x)$ απαιτώντας η συνάρτηση συμμετοχής του να έχει τη μορφή A . αν δηλαδή υποθέσουμε ότι η συνάρτηση συμμετοχής είναι η κανονική, τότε για να υλοποιηθεί πρέπει η συνάρτηση ενεργοποίησης να είναι :

$$\alpha(u) = \exp\left(-\frac{(u-m)^2}{\sigma^2}\right), \text{ όπου } m \text{ το κέντρο της συνάρτησης συμμετοχής και } \sigma \text{ το εύρος της.}$$

ΛΕΙΤΟΥΡΓΙΚΟΙ ΝΕΥΡΩΝΕΣ

Με τον ίδιο τρόπο μπορούν να υλοποιηθούν οι βασικές πράξεις των ασαφών συνόλων. Οι πράξεις αυτές στηρίζονται σε λειτουργίες όπως η τ - νόρμα η σ - νόρμα, το ασαφές συμπλήρωμα κλπ. Οι παραπάνω λειτουργίες μπορούν να αντικαταστήσουν τη συνάρτηση ενεργοποίησης του νευρώνα αν θεωρήσουμε ότι οι είσοδοι δεν αθροίζονται πολλαπλασιασμένοι με βάρη, όπως στους κλασσικούς νευρώνες. Για να μπορούμε όμως να εφαρμόσουμε τους αλγόριθμους μάθησης (πχ τον backpropagation αλγόριθμος αντίστροφης διάδοσης) θα πρέπει οι λειτουργίες να είναι παραγωγίσιμες.

Για την υλοποίηση των ασαφών λειτουργιών μπορούμε να επεκτείνουμε την έννοια της συνάρτησης με ταφοράς του νευρώνα. Στην περίπτωση αυτή ο νευρώνας δεν υλοποιεί το αναλυτικό γινόμενο της εισόδου με το διάνυσμα των βαρών αλλά την πράξη της συγκεκριμένης ασαφούς λειτουργίας θεωρώντας ότι όλα τα βάρη είναι 1. Ως συνάρτηση συμμετοχής λαμβάνεται η :

$$\alpha(u) = \begin{cases} 1, & u > 1 \\ u, & 0 \leq u \leq 1 \\ 0, & u < 0 \end{cases}$$

ΣΥΝΘΕΤΙΚΟΙ ΝΕΥΡΩΝΕΣ

Η πράξη της σύνθεσης αποτελεί μια γενίκευση της πράξης του αναλυτικού γινομένου των διανυσμάτων και υλοποιείται από τη συνάρτηση μεταφοράς f των νευρώνων. Η μορφή του νευρώνα είναι η ίδια και ορίζεται από τις εξισώσεις:

$$f = \text{union}_{t \in \{1, \dots, m\}} (x_t, w_t)$$

$$y = a(u)$$

Όπου union είναι μια σ -νόρμα και t είναι μια τ -νόρμα. Ως συνάρτηση ενεργοποίησης θεωρείται η συνάρτηση αναρρίχησης. Παρατηρούμε ότι ο συνθετικός νευρώνας είναι μια γενίκευση του κλασσικού νευρώνα, αφού η πράξη της φραγμένης άθροισης, η οποία εξασφαλίζεται από τη συνάρτηση ενεργοποίησης είναι μια ειδική περίπτωση της φραγμένης ένωσης, ενώ το γινόμενο είναι μια ειδική περίπτωση της ασαφούς τομής.

Προσαρμοστικά Δίκτυα (Adaptive Networks)

Ένα προσαρμοστικό δίκτυο είναι αυτό που στη δομή του περιέχει έναν αριθμό κόμβων συνδεδεμένων μέσω κατευθυντικών συνδέσμων. Κάθε κόμβος αναπαριστά μια μονάδα επεξεργασίας και οι σύνδεσμοι μεταξύ κόμβων προσδιορίζουν την αιτιολογική σχέση (causal relationship) μεταξύ των συνδεδεμένων κόμβων. Προσαρμοστικός είναι ο κόμβος, όλος ή μέρος του, που σημαίνει ότι οι εξοδοί αυτών των κόμβων εξαρτώνται από προσαρμοζόμενες – τροποποιήσιμες παραμέτρους που ανήκουν σε αυτούς τους κόμβους. Οι κανόνες εκπαίδευσης προσδιορίζουν πως αυτές οι παράμετροι θα πρέπει να αλλάζουν ώστε να ελαχιστοποιείται το προκαθορισμένο μέτρο σφάλματος (error measure), το οποίο είναι μια μαθηματική έκφραση που μετράει την ασυμφωνία μεταξύ της πραγματικής εξόδου του δικτύου και της επιθυμητής εξόδου.

Τα προσαρμοστικά δίκτυα χρησιμοποιούνται στην ταυτοποίηση συστήματος. Εμείς θα πρέπει να βρούμε μια κατάλληλη αρχιτεκτονική για το δίκτυο και να θέσουμε ένα σύνολο παραμέτρων που μπορούν να μοντελοποιήσουν καλύτερα ένα σύστημα -στόχο, το οποίο περιγράφεται από ένα σύνολο ζευγών δεδομένων εισόδου -εξόδου. Ο βασικός κανόνας ενός προσαρμοστικού δικτύου είναι η απότομη φθίνουσα μέθοδος, στην οποία το βαθμωτό διάνυμα προκύπτει από επιτυχείς επικλήσεις του κανόνα της αλυσίδας. Αυτή η μέθοδος χρησιμοποιείται και για την εύρεση του βαθμωτού σε ένα νευρωνικό δίκτυο πολλών επιπέδων. Αυτή η μέθοδος ονομάζεται κανόνας οπισθοδρόμησης διάδοσης (back propagation learning rule).

ΠΡΟΣΑΡΜΟΣΤΙΚΟ ΝΕΥΡΟ-ΑΣΑΦΕΣ ΣΥΣΤΗΜΑ ΕΞΑΓΩΓΗΣ ΣΥΜΠΕΡΑΣΜΑΤΟΣ (ADAPTIVE NEURO-FUZZY INFERENCE SYSTEMS-ANFIS).

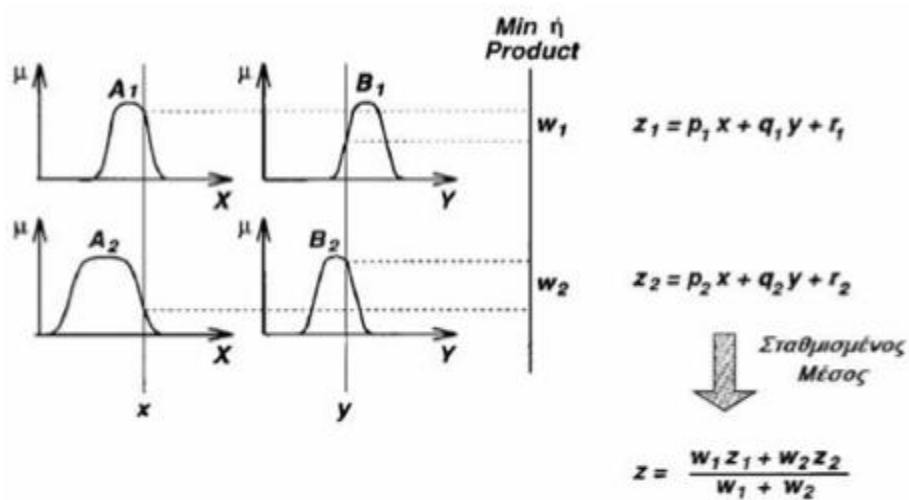
Η Αρχιτεκτονική του ANFIS

Ο αλγόριθμος ANFIS είναι από τους κυριότερους αλλά και ταυτόχρονα από τους πρώτους που εφαρμόστηκαν στο πεδίο της νευρο-ασαφούς προσέγγισης προβλημάτων. Παρακάτω περιγράφεται το δίκτυο, υποθέτοντας ότι το πρόβλημα το οποίο θα αναλυθεί έχει δύο εισόδους x και y και μία έξοδο z . Υποθέτοντας ότι για ένα πρώτης τάξης μοντέλο Sugeno, μία τυπική βάση κανόνων (rule base) θα μπορούσε να είναι και η εξής:

Κανόνας 1: Εάν x είναι A_1 και y είναι B_1 τότε $f_1 = p_1 \times x + q_1 \times y + r_1$

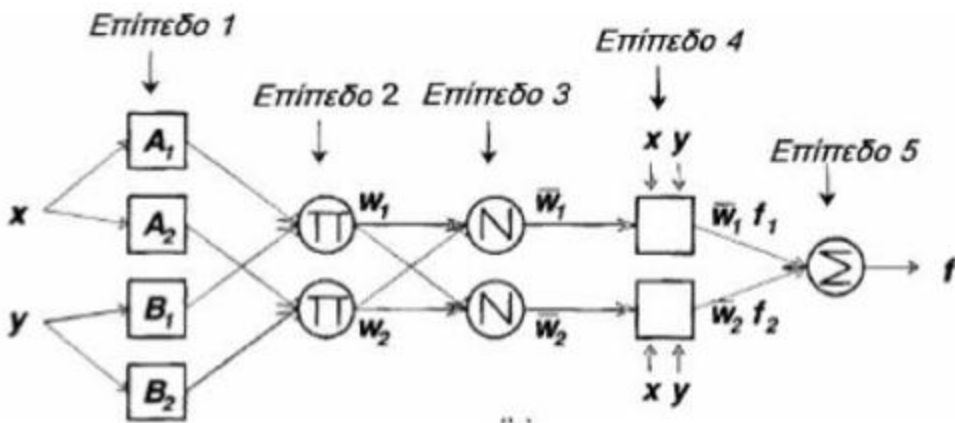
Κανόνας 2: Εάν x είναι A_2 και y είναι B_2 τότε $f_2 = p_2 \times x + q_2 \times y + r_2$

Το παρακάτω σχήμα δείχνει με απλό τρόπο τη διαδικασία συμπερασμού (inference procedure) του μοντέλου Sugeno, στην περίπτωση όπου για t- operator έχει επιλεγεί η τομή των δύο ασαφών συνόλων (A,B)



Σχήμα :Ασαφής συλλογιστική

Όπως γίνεται φανερό, η έξοδος z του πρωτοβάθμιου μοντέλου Sugeno είναι ένας σταθμικός μέσος όρος. Η αντίστοιχη αναπαράσταση του δικτύου ANFIS παρουσιάζεται στο παρακάτω σχήμα



Σχήμα : Η αρχιτεκτονική του ANFIS

Το παραπάνω σχήμα απεικονίζει τον συλ λογιστικό μηχανισμό (reasoning) για αυτό το μοντέλο Sugeno και η αντίστοιχη ισοδύναμη αρχιτεκτονική του ANFIS όπου οι κόμβοι του ίδιου επιπέδου έχουν παρόμοιες συναρτήσεις.

Περιορισμοί του ANFIS

Οι κυριότεροι περιορισμοί του ANFIS που αναφέρονται στην διεθνή βιβλιογραφία είναι:

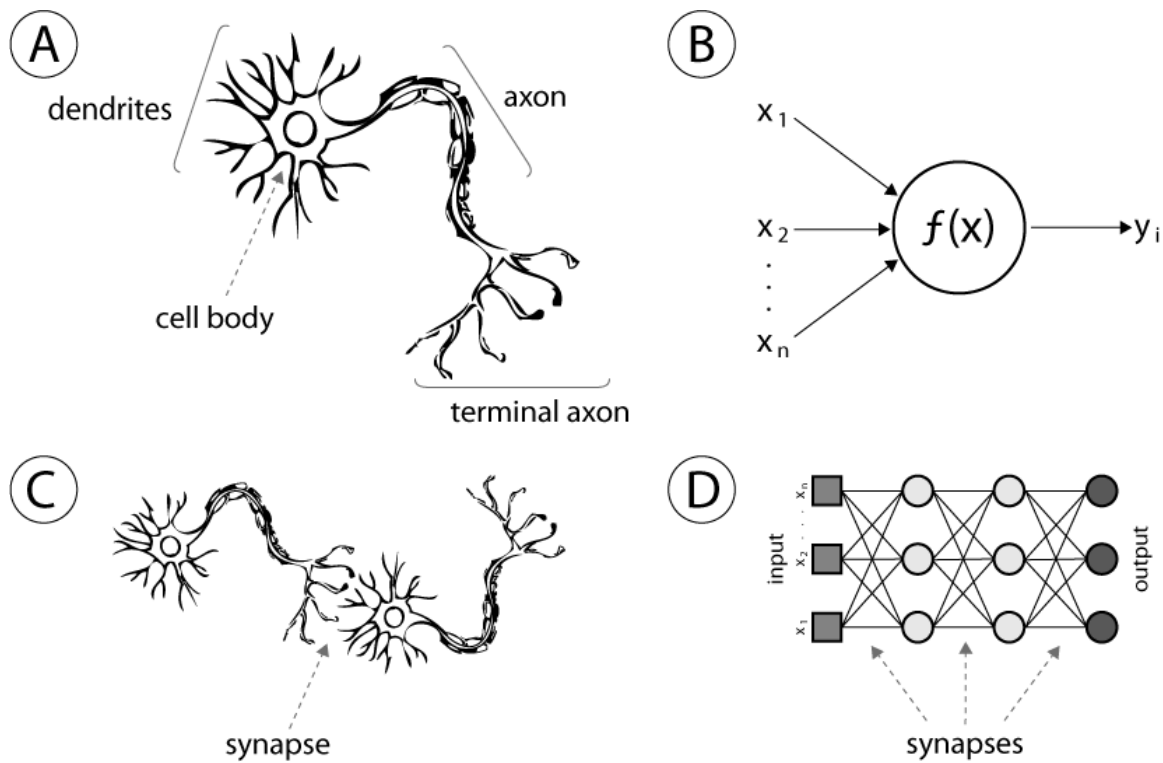
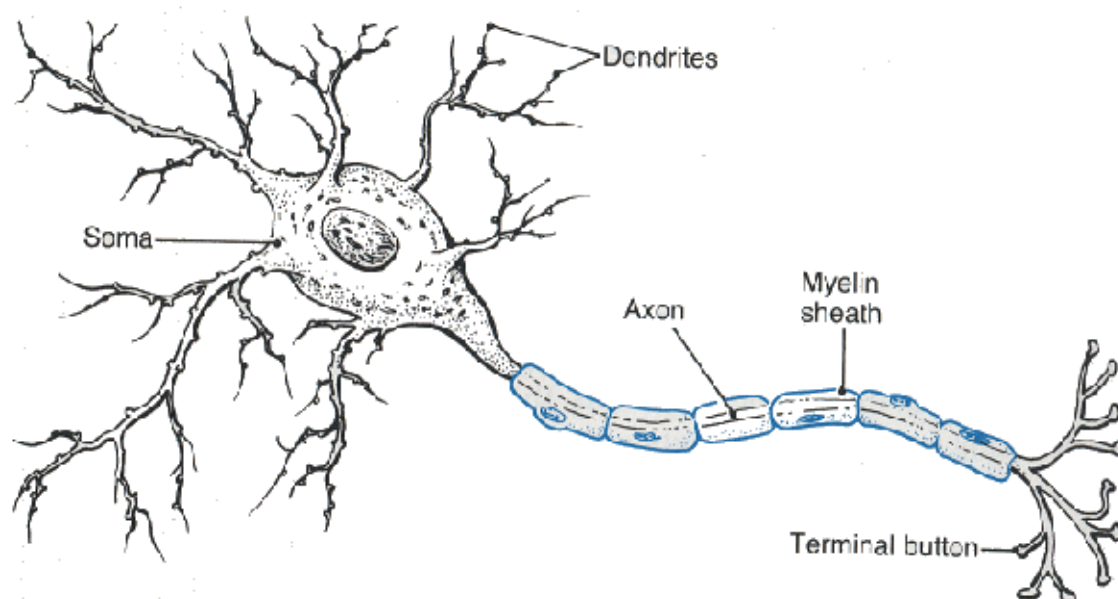
1. Βασίζεται σε ένα FIS τύπου Sugeno.
2. Έχει μια έξοδο, που λαμβάνεται με τη μέθοδο αποασαφοποίησης
3. σταθμισμένου μέσου. Όλες οι συναρτήσεις συμμετοχής εξόδου πρέπει να είναι ίδιου τύπου, είτε γραμμικές είτε σταθερές.
4. Δεν μπορεί να γίνεται κοινή χρήση κανόνων. Διαφορετικοί κανόνες δεν μπορούν να έχουν την ίδια συνάρτηση συμμετοχής εξόδου.

5. Πρέπει να υπάρχουν βάρη σε κάθε κανόνα.

6. Δεν μπορεί να χρησιμοποιηθεί οποιαδήποτε συνάρτηση συμμετοχής, παρά μόνο προκαθορισμένες που επιβάλλουν οι περιορισμοί του ANFIS.

Τεχνητά Νευρωνικά Δίκτυα

Τα Τεχνητά Νευρωνικά Δίκτυα αποτελούν μια προσπάθεια προσέγγισης της λειτουργίας του ανθρώπινου εγκεφάλου από μια μηχανή. Έχουν την ικανότητα να εκτελούν υπολογισμούς με μαζικό παράλληλο τρόπο. Η αρχιτεκτονική τους βασίζεται στην αρχιτεκτονική των Βιολογικών Δικτύων. Τα Τεχνητά Νευρωνικά Δίκτυα είναι μια συλλογή από νευρώνες (Processing Units-PUs) που συνδέονται μεταξύ τους. Κάθε νευρώνας έχει πολλές εισόδους αλλά μόνο μία έξοδο η οποία με τη σειρά της μπορεί να αποτελέσει είσοδο για επόμενους νευρώνες. Οι συνδέσεις μεταξύ των νευρώνων διαφέρουν ως προς τη σημαντικότητά τους, η οποία και προσδιορίζεται από το συντελεστή βαρύτητας. Η επεξεργασία κάθε νευρώνα καθορίζεται από τη συνάρτηση μεταφοράς, η οποία καθορίζει την κάθε έξοδο σε σχέση με τις εισόδους και τους συντελεστές βάρους.



Σχήμα : παρατηρούμε την παρόμοια αρχιτεκτονική μεταξύ του βιολογικού νευρώνα και του τεχνικού νευρωνικού δικτύου.

Τα νευρωνικά δίκτυα (neural networks ή με σύντμηση neural nets) αποτελούν μια σχετικά νέα περιοχή στις φυσικές επιστήμες, καθόσον έχουν γίνει γνωστά και έχουν αναπτυχθεί σε διεθνές επίπεδο μόνο κατά τις τελευταίες δεκαετίες. Εν τούτοις, η περιοχή αυτή έχει δει μια μεγάλη άνθηση, η οποία διαφαίνεται από την μεγάλη ανάπτυξη που έχει παρατηρηθεί, από τον αριθμό των επιστημόνων που ασχολούνται με αυτά τα θέματα και βέβαια από τα πολύ σημαντικά επιτεύγματα που έχουν συμβάλλει στο να γίνουν τα νευρωνικά δίκτυα γνωστά σε ένα ευρύτερο κύκλο. Αποτελούν επομένως ένα θέμα με μεγάλο ενδιαφέρον στις τεχνολογικές επιστήμες. Το κύριο χαρακτηριστικό τους είναι ότι οι πρώτες αρχές και λειτουργίες τους βασίζονται και εμπνέονται από το νευρικό σύστημα των ζώντων οργανισμών (και φυσικά του ανθρώπου), αλλά η μελέτη και η χρήση τους έχει προχωρήσει πολύ πέρα από τους βιολογικούς οργανισμούς. Ουσιαστικά δημιουργήθηκε μία νέα περιοχή η οποία έχει αποκοπεί τελείως από την βιολογία και σήμερα τα νευρωνικά δίκτυα χρησιμοποιούνται για να λύσουν κάθε είδους προβλήματα με ηλεκτρονικό υπολογιστή. Η φιλοσοφία τους όμως είναι διαφορετική από τον τρόπο με τον οποίο δουλεύουν οι κλασικοί υπολογιστές. Η λειτουργία τους προσπαθεί να συνδυάσει τον τρόπο σκέψης του ανθρώπινου εγκεφάλου με τον αφηρημένο μαθηματικό τρόπο σκέψης. Έτσι στα νευρωνικά δίκτυα χρησιμοποιούμε τέτοιες ιδέες όπως, λ.χ. ένα δίκτυο μαθαίνει και εκπαιδεύεται, θυμάται ή ξεχνά μια αριθμητική τιμή κτλ., πράγματα που μέχρι τώρα τα αποδίδαμε μόνο στην ανθρώπινη σκέψη. Αλλά βέβαια μπορούν και χρησιμοποιούν επί πλέον και περίπλοκες μαθηματικές συναρτήσεις και κάθε είδους εργαλεία από την μαθηματική ανάλυση.

Ένα ιδιαίτερο χαρακτηριστικό είναι ότι οι επιστήμονες στην περιοχή των νευρωνικών δικτύων προέρχονται σχεδόν από όλες τις περιοχές των φυσικών επιστημών, όπως την Ιατρική, την επιστήμη Μηχανικών, τη Φυσική, τη Χημεία, τα Μαθηματικά, την επιστήμη Υπολογιστών, την

Ηλεκτρολογία κτλ. Αυτό δείχνει ότι για την ανάπτυξή τους απαιτούνται ταυτόχρονα γνώσεις και θέματα από πολλές περιοχές, ενώ το ίδιο ισχύει και για τις τεχνικές και τις μεθόδους που χρησιμοποιούνται. Έτσι καταλαβαίνει κανείς ότι τα νευρωνικά δίκτυα δίνουν μια νέα πρόκληση στις επιστήμες, καθόσον οι νέες γνώσεις που απαιτούνται είναι από τις πιο χρήσιμες στον άνθρωπο, τόσο για την ζωή και την ιατρική όσο και για την τεχνολογία. Καμία άλλη επιστήμη σήμερα δεν συνδυάζει με τόσο άμεσο τρόπο γνώσεις που προέρχονται από τόσο διαφορετικές περιοχές.

Η έμπνευση για τα νευρωνικά δίκτυα, όπως αναφέρθηκε παραπάνω, ξεκινά από την βιολογία. Οι ζώντες οργανισμοί, από τους πιο απλούς μέχρι τον άνθρωπο, έχουν ένα νευρικό σύστημα το οποίο είναι υπεύθυνο για μια πλειάδα από διεργασίες, όπως είναι η επαφή με τον εξωτερικό κόσμο, η μάθηση, η μνήμη κτλ. Το νευρικό σύστημα των οργανισμών αποτελείται από πολλά νευρωνικά δίκτυα τα οποία είναι εξειδικευμένα στις διεργασίες αυτές. Η κεντρική μονάδα του νευρικού συστήματος είναι, οπωσδήποτε, ο εγκέφαλος, ο οποίος επίσης αποτελείται από νευρωνικά δίκτυα. Κάθε νευρωνικό δίκτυο αποτελείται από ένα μεγάλο αριθμό μονάδων, που λέγονται νευρώνες ή νευρώνια (neurons). Ο νευρώνας είναι η πιο μικρή ανεξάρτητη μονάδα του δικτύου, όπως λ.χ. το άτομο είναι η πιο μικρή μονάδα της ύλης. Οι νευρώνες συνεχώς και ασταμάτητα επεξεργάζονται πληροφορίες, παίρνοντας και στέλνοντας ηλεκτρικά σήματα σε άλλους νευρώνες.

Βλέπουμε λοιπόν ότι οι πρώτες γνώσεις μας για τα νευρωνικά δίκτυα προέρχονται από την βιολογία και την ιατρική. Σήμερα διεξάγεται ιδιαίτερα μεγάλη έρευνα στις δύο αυτές επιστήμες για την καλύτερη κατανόηση των νευρωνικών δικτύων του εγκεφάλου, καθόσον είναι προφανές ότι αυτό θα βοηθήσει στο να εξηγήσουμε πώς ακριβώς λειτουργεί ο εγκέφαλος και τις τόσο

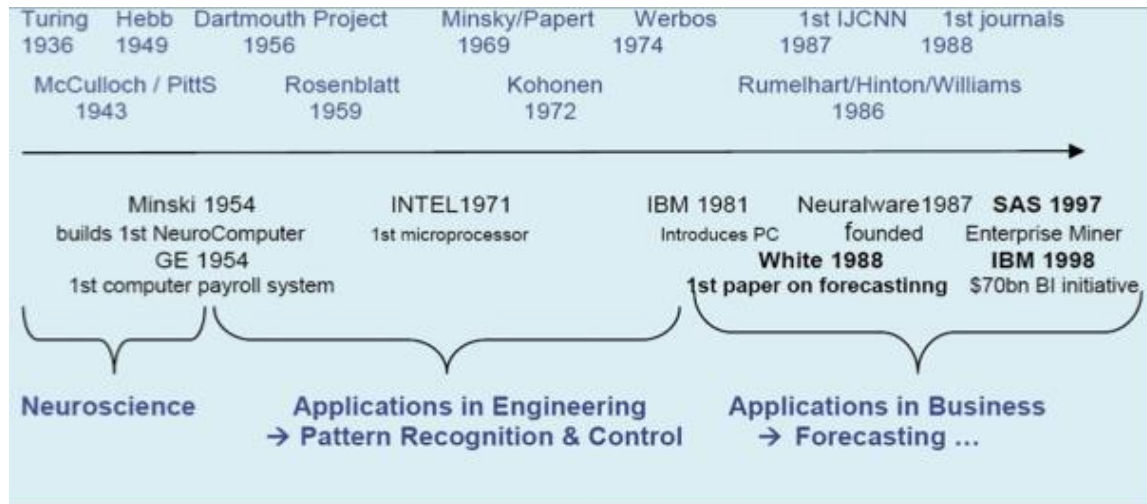
περίπλοκες διεργασίες του, όπως πως σκεπτόμαστε, πως θυμόμαστε κτλ. Οι έννοιες αυτές, παρόλο ότι ακούγονται απλές, εν τούτοις δεν έχουν εξηγηθεί σχεδόν καθόλου μέχρι σήμερα από τους επιστήμονες. Έτσι, λοιπόν, τα νευρωνικά δίκτυα των ζώντων οργανισμών τα ονομάζουμε βιολογικά νευρωνικά δίκτυα, ενθυμούμενοι ότι αυτά είναι και τα πρώτα δίκτυα που μελετήθηκαν, καθώςον υπάρχουν σε όλους τους ζώντες οργανισμούς (όχι όμως στα φυτά).

Οι διεργασίες που επιτελούνται από τα βιολογικά νευρωνικά δίκτυα στους ζώντες οργανισμούς είναι πολύ περίπλοκες αλλά και τόσο χρήσιμες στην καθημερινή ζωή του ανθρώπου. Η ανάπτυξη της τεχνολογίας και η επίλυση αναγκών έχει οδηγήσει στο να γίνουν κάποιες πρώτες σκέψεις μήπως είναι δυνατόν να δημιουργηθούν κάποια πρότυπα (μοντέλα) του νευρωνικού συστήματος του ανθρώπου, τα οποία θα περιέχουν όλα τα χαρακτηριστικά που είναι γνωστά μέχρι σήμερα και τα οποία θα μπορούσαν από μόνα τους να επιτελέσουν τις εργασίες αυτές, με τον ίδιο τρόπο που γίνονται στα βιολογικά νευρωνικά δίκτυα. Τα δίκτυα αυτά ονομάζονται τεχνητά νευρωνικά δίκτυα (artificial neural nets, ANN). Η βασική τους διαφορά από τα βιολογικά δίκτυα είναι ότι τα δίκτυα αυτά παίρνουν γνώσεις (μαθαίνουν) με την εξάσκηση και την εμπειρία, όπως ακριβώς και οι άνθρωποι, αλλά διαφέρουν στο ότι δεν ακολουθούν ορισμένους προκαθορισμένους κανόνες, που είναι χαρακτηριστικό των υπολογιστών. Υπάρχει σήμερα ένας μεγάλος όγκος έρευνας στην περιοχή αυτή, καθώς και εδώ είναι προφανές πόσο χρήσιμο θα ήταν να μπορεί κάτι το άψυχο να επιτελεί εργασίες που μέχρι σήμερα μόνο ο άνθρωπος μπορούσε να κάνει, είτε αυτό είναι μια μηχανή είτε ένα πρόγραμμα ηλεκτρονικού υπολογιστή.

ΙΣΤΟΡΙΚΗ ΕΞΕΛΙΞΗ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ

Η έρευνα γύρω από τα ΝΔ. έχει περάσει περιόδους έντονης δραστηριότητας. Η αρχή έγινε την δεκαετία του 1940 με την εργασία των McCulloch και Pits (1943). Το 1960 ο Rosenblatt διατύπωσε το θεώρημα της σύγκλισης για τα Perceptron. Η πρώτη απόπειρα εφαρμογής των νευρωνικών δικτύων σε πραγματικά δεδομένα έγινε το 1964. Ο Hu (1964) στην έρευνά του χρησιμοποιεί το προσαρμοζόμενο γραμμικό μοντέλο του Widrow για την πρόβλεψη των καιρικών συνθηκών. Εξαιτίας της τότε έλλειψης αλγορίθμου εκπαίδευσης για πολύστρωματικά δίκτυα, η έρευνα ήταν αρκετά περιορισμένη. Μεγάλη εξέλιξη στην χρησιμοποίηση των ΝΔ για την παραγωγή προβλέψεων πραγματοποιήθηκε με την ανάπτυξη του αλγορίθμου backpropagation (Rumelhart et al., 1986b). Ο Werbos (1974),(1988) πρώτος εφάρμοσε τον αλγόριθμο backpropagation και βρήκε ότι τα ΝΔ που χρησιμοποίησαν τον αλγόριθμο αυτό είχαν καλύτερα αποτελέσματα από τις παραδοσιακές στατιστικές μεθόδους όπως η απλή γραμμική παλινδρόμηση και οι προσεγγίσεις των Box- Jenkins. Οι Lapedes και Farber (1987) διεξήγαγαν μια εμπειρική μελέτη και κατέληξαν ότι τα ΝΔ μπορούν να χρησιμοποιηθούν για την μοντελοποίηση και την πρόβλεψη μη γραμμικών χρονοσειρών. Οι Weigend et al.(1990), Cottrell et al.(1995) έφεραν στην επιφάνεια το θέμα της αρχιτεκτονικής του δικτύου όταν χρησιμοποιείται για πρόβλεψη πραγματικών χρονοσειρών. Οι Tang et al.(1991), Sharda, Patil (1992) και Tang, Fishwick(1993) , μεταξύ άλλων, παρουσίασαν αποτελέσματα από διάφορες συγκρίσεις προβλέψεων μεταξύ μεθόδου Box-Jenkins και μοντέλων ΝΔ. Σε ένα διαγωνισμό προβλέψεων που διοργανώθηκε το 1993 από τους Weigend και Gershenfeld σε συνεργασία με το Santa Fe Institute, οι νικητές σε κάθε σύνολο δεδομένων χρησιμοποίησαν νευρωνικά δίκτυα (Gershenfeld και Weigend,1993). Αξιοσημείωτες είναι οι ερευνητικές προσπάθειες στον τομέα

των προβλέψεων με χρήση Νευρωνικών Δικτύων. Οι Marquez et al.(1992) και Hill et al.(1994) επαναλαμβάνοντας την ήδη υπάρχουσα γνώση για τα ΝΔ, συγκρίνανε τα ΝΔ με στατιστικά μοντέλα των μεθόδων χρονοσειρών και των μεθόδων παλινδρόμησης. Ακολουθεί ο **πίνακας** με



τα κυριότερα βήματα στην εξέλιξη και εφαρμογή των ΝΔ.

Πίνακας : Τα κυριότερα στάδια εξέλιξης των νευρωνικών δικτύων

Εκπαίδευση νευρωνικού δικτύου

Για να χρησιμοποιηθεί ένα Νευρωνικό Δίκτυο πρέπει πρώτα να εκπαιδευτεί για να μάθει. Η μάθηση συνίσταται στον προσδιορισμό των κατάλληλων συντελεστών βάρους, ώστε το Νευρωνικό Δίκτυο να εκτελεί τους επιθυμητούς υπολογισμούς, και πραγματοποιείται με τη βοήθεια αλγόριθμων που είναι γνωστοί ως κανόνες μάθησης (training). Ο ρόλος των συντελεστών βάρους μπορεί να ερμηνευτεί ως αποθήκευση γνώσης, η οποία παρέχεται μέσω παραδειγμάτων. Ανάμεσα στις πολλές ενδιαφέρουσες ιδιότητες ενός νευρωνικού δικτύου αυτή με τη μεγαλύτερη σπουδαιότητα είναι η ικανότητα του να εκπαιδεύεται από το περιβάλλον του και έτσι να βελτιώνει την απόδοσή του μέσω της εκπαίδευσης. Η βελτίωση αυτή γίνεται σταδιακά, με το χρόνο, σύμφωνα με κάποιο καθορισμένο μέτρο. Η εκπαίδευση επιτυγχάνεται μέσω μιας επαναληπτικής διαδικασίας ρυθμίσεων της τιμής των βαρών. Θεωρητικά, το δίκτυο αποκτά περισσότερη γνώση για το περιβάλλον του μετά από κάθε επανάληψη της διαδικασίας εκπαίδευσης.

Για να ορίσουμε μια έννοια όπως αυτή της εκπαίδευσης εξαρτάται από ποια σκοπιά θα την εξετάσουμε. Εμείς με σημείο αναφοράς τα νευρωνικά δίκτυα χρησιμοποιούμε τον ορισμό των Mendel και McClaren (1970):

Εκπαίδευση είναι μια διαδικασία με την οποία προσαρμόζονται οι ελεύθερες παράμετροι ενός νευρωνικού δικτύου μέσω μίας συνεχούς διαδικασίας διέγερσης από το περιβάλλον στο οποίο βρίσκεται το δίκτυο. Το είδος της μάθησης καθορίζεται από τον τρόπο με τον οποίο πραγματοποιούνται οι αλλαγές των παραμέτρων.

Ο παραπάνω ορισμός της διαδικασίας μάθησης υπονοεί την ακόλουθη σειρά βημάτων:

1. Το νευρωνικό δίκτυο "διεγείρεται" από ένα περιβάλλον.

2. Το νευρωνικό δίκτυο υφίσταται αλλαγές σαν συνέπεια αυτής της διέγερσης.
3. Το νευρωνικό δίκτυο αλληλεπιδρά (feedback) με ένα καινούργιο τρόπο στο περιβάλλον, λόγω των αλλαγών που συνέβησαν στην εσωτερική του δομή.

Πιο συγκεκριμένα ας αναλύσουμε την παραπάνω περιγραφή με ένα παράδειγμα. Θεωρούμε ένα ζεύγος κομβικά σήματα x_j και u_k που συνδέονται με βάρος σύνδεσης w_{kj} . Το x_j αναπαριστά την είσοδο του νευρώνα j , ενώ το u_k την εσωτερική λειτουργία του νευρώνα k . Με βάση το βάρος σύνδεσης w_{kj} , τα

σήματα x_j και u_k αναφέρονται συχνά σαν προσυναπτική και μετασυναπτική λειτουργία αντίστοιχα. Ας συμβολίσουμε με $w_{kj}(n)$ την τιμή του βάρους σύνδεσης w_{kj} την χρονική στιγμή n . Τη χρονική αυτή στιγμή γίνεται μια ρύθμιση (διόρθωση) $\Delta w_{kj}(n)$ στο βάρος $w_{kj}(n)$ και παράγεται η νέα ενημερωμένη τιμή $w_{kj}(n+1)$. Έτσι προκύπτει η εξίσωση :

$$w_{kj}(n+1) = w_{kj}(n) + \Delta w_{kj}(n)$$

όπου $w_{kj}(n)$ και $w_{kj}(n+1)$ η παλιά και η νέα τιμή του βάρους σύνδεσης w_{kj} , αντίστοιχα. Η εξίσωση συγκεντρώνει την συνολική επίδραση των βημάτων 1 και 2 που προκύπτουν από τον ορισμό της διαδικασίας μάθησης και αναφέρθηκαν παραπάνω. Πιο συγκεκριμένα η ρύθμιση $\Delta w_{kj}(n)$ υπολογίζεται σαν αποτέλεσμα της διέγερσης από το περιβάλλον (βήμα 1) και η ενημερωμένη τιμή $w_{kj}(n+1)$ εκφράζει την αλλαγή που συνέβη στο δίκτυο σαν αποτέλεσμα αυτής της διέγερσης (βήμα 2). Το βήμα 3 πραγματοποιείται όταν υπολογίζεται η απάντηση του νέου δικτύου που λειτουργεί με το ενημερωμένο σύνολο παραμέτρων $\{w_{kj}(n+1)\}$.

Ένα καθορισμένο σύνολο από καλά ορισμένους κανόνες για τη λύση ενός προβλήματος μάθησης καλείται αλγόριθμος εκπαίδευσης (learning algorithm).

Όπως είναι φανερό, δεν υπάρχει ένας μοναδικός τέτοιος αλγόριθμος για το σχεδιασμό νευρωνικών δικτύων. Αντίθετα, υπάρχει ένα σύνολο από εργαλεία που αναπαρίστανται από μια μεγάλη ποικιλία αλγορίθμων εκπαίδευσης, καθένας από τους οποίους έχει τα δικά του πλεονεκτήματα. Μερικοί από τους πιο γνωστούς αλγόριθμους εκπαίδευσης είναι οι ακόλουθοι :

Η μέθοδος Ελαχίστου Μέσου Τετραγωνικού Λάθους (LSE)

Η μέθοδος Ελάχιστου Μέσου Τετραγωνικού Λάθους (Least Squares Error) βασίζεται στον αλγόριθμο Ελαχίστων Μέσων Τετραγώνων (Least Mean Square – LMS), που είναι επίσης γνωστός σαν Delta-rule ή σαν ο κανόνας των Widrow και Hoff (1960). Ο αλγόριθμος LMS λειτουργεί με το μοντέλο ενός απλού γραμμικού νευρώνα, και έχει βρει πολλές εφαρμογές. Πράγματι, ο LMS αλγόριθμος καθιερώθηκε σαν ένα σπουδαίο λειτουργικό κομμάτι στην συνεχώς επεκτεινόμενη περιοχή της προσαρμοζόμενης επεξεργασίας σημάτων. Ως πλεονεκτήματα του αλγορίθμου LMS, μπορούμε να αναφέρουμε τα εξής :

- 1) Υπολογίζει εκτιμήσεις των χωρικών συναρτήσεων συσχέτισης με ένα απλό και συγχρόνως αποδοτικό τρόπο
- 2) Ελαχιστοποιεί το στιγμιαίο λάθος $E(n)$, άρα ελαχιστοποιεί τις απαιτήσεις μνήμης.
- 3) Λειτουργεί τόσο σε στάσιμο όσο και σε μη-στάσιμο περιβάλλον

Η μέθοδος Ταχύτερης Καθόδου (Steepest Descent)

Σύμφωνα με αυτή τη μέθοδο, υποθέτουμε ότι τα βάρη του φίλτρου είναι χρονικά μεταβαλλόμενα και ότι οι τιμές τους διορθώνονται με ένα επαναληπτικό τρόπο κατά μήκος της επιφάνειας λάθους, μετακινώντας τα προοδευτικά προς τη βέλτιστη λύση. Η μέθοδος ταχύτερης καθόδου έχει σαν στόχο τη συνεχή αναζήτηση βέλτιστης λύσης. Ανεξάρτητα από ποια προσέγγιση θα χρησιμοποιήσουμε, για να δουλέψει η μέθοδος ταχύτερης καθόδου, πρέπει να δώσουμε ιδιαίτερη προσοχή στην επιλογή της παραμέτρου μάθησης. Η μέθοδος αυτή δεν χρησιμοποιείται συχνά λόγω των μειονεκτημάτων τα οποία είναι τα εξής :

Για να δουλέψει αυτή η μέθοδος πρέπει να δώσουμε ιδιαίτερη προσοχή στην επιλογή της παραμέτρου μάθησης. Απαιτεί τη γνώση των χωρικών συναρτήσεων συσχέτισης, οι οποίες είναι συνήθως άγνωστες. Ελαχιστοποιεί το άθροισμα των τετραγώνων του λάθους, για όλες τις επαναλήψεις, άρα απαιτεί την αποθήκευση μεγαλύτερου όγκου πληροφορίας.

Ο αλγόριθμος Οπισθόδρομης Μάθησης (Backpropagation Learning Algorithm)

Είναι ο πιο διαδεδομένος αλγόριθμος εκμάθησης και βασίζεται στον κανόνα μάθησης και διόρθωσης του λάθους (error correction learning rule). Βασικά η διαδικασία της Οπισθόδρομης Μάθησης αποτελείται από δυο περάσματα διαμέσου των διαφορετικών επιπέδων του δικτύου ένα προς τα εμπρός πέρασμα (forward pass) και ένα προς τα πίσω πέρασμα (backward pass).

Στο εμπρός πέρασμα ένα διάνυσμα εισόδου (input vector) εφαρμόζεται στους νευρώνες εισόδου του δικτύου, και η επίδραση του διαδίδεται μέσα στο δίκτυο από επίπεδο σε επίπεδο (layer by

layer). Τελικά ένα σύνολο από εξόδους παράγεται ως η πραγματική απόκριση του δικτύου. Κατά τη διάρκεια του εμπρός περάσματος τα βάρη του δικτύου είναι σταθερά.

Από την άλλη μεριά κατά τη διάρκεια της πίσω διάδοσης τα βάρη προσαρμόζονται σε συμφωνία με τον κανόνα διόρθωσης λάθους. Πιο συγκεκριμένα, η πραγματική απόκριση του δικτύου αφαιρείται από την επιθυμητή απόκριση για την παραγωγή ενός σήματος λάθους, που διαδίδεται προς τα πίσω στο δίκτυο, αντίθετα από την κατεύθυνση των συνδέσεων, από το οποίο προκύπτει και το όνομα πίσω διάδοσης του λάθους.

Τα συναπτικά βάρη προσαρμόζονται έτσι ώστε να κάνουν την πραγματική απόκριση του δικτύου να πλησιάσει την επιθυμητή απόκριση.

Στην βιβλιογραφία ο αλγόριθμος Οπισθόδρομης Μάθησης συχνά αναφέρεται και σαν Back Propagation Algorithm ή πιο απλά σαν Back Prop. Από δω και στο εξής θα αναφερόμαστε σε αυτόν σαν αλγόριθμο πίσω διάδοσης ή Π.Δ.. Η διαδικασία μάθησης που εκτελείται με αυτόν τον αλγόριθμο ονομάζεται μάθηση πίσω διάδοσης.

Αρχιτεκτονικές Τεχνητών Νευρωνικών Δικτύων

Ο τρόπος με τον οποίο οι νευρώνες ενός νευρωνικού δικτύου δομούνται, είναι στενά συνδεδεμένος με τον αλγόριθμο εκπαίδευσης που χρησιμοποιείται για την εκπαίδευση του δικτύου. Σαν εμπρόσθιας τροφοδότησης (FeedForward), αναφέρονται τα δίκτυα, στα οποία τα σήματα κατευθύνονται από την είσοδο στην έξοδο. Όταν οι εξοδοί κάποιων νευρώνων, γίνονται είσοδοι σε νευρώνες προηγούμενων επιπέδων (προς το μέρος της εισόδου του δικτύου), τότε έχουμε ανάδραση. Οι πιο διαδεδομένες αρχιτεκτονικές τεχνητών νευρωνικών δικτύων είναι επιγραμματικά οι εξής:

- **Νευρωνικό Δίκτυο Οπίσθιας Τροφοδότησης (Back Propagation Neural Network)**

Ο αλγόριθμος όπισθεν διάδοσης σφαλμάτων είναι μια γενική μέθοδος για υπολογισμό των βαρών ενός πολυστρωματικού δικτύου. Στόχος της μεθόδου είναι να εκπαιδευτούν τα βάρη με βάση το σύνολο των τιμών των ζευγών εισόδου-εξόδου,

καθώς το δίκτυο ανήκει στη κατηγορία supervised learning, μειώνοντας με κάθε εποχή λειτουργίας του αλγορίθμου το σήμα σφάλματος στην έξοδο του νευρωνικού δικτύου

- **Νευρωνικό Δίκτυο Εμπρόσθιας Τροφοδότησης (FeedForward Neural Network)** του οποίου η δομή είναι 3-3-1. Δηλαδή αποτελείται από ένα επίπεδο εισόδου (input layer) με τρεις νευρώνες, ένα κρυμμένο επίπεδο (hidden layer) με τρεις νευρώνες και ένα επίπεδο εξόδου (output layer) με μια και μοναδική έξοδο που είναι η πρόβλεψη για την επόμενη ημέρα. Η κάθε είσοδος του δικτύου αντιστοιχίζεται με όλους τους νευρώνες του κρυμμένου επιπέδου όπου κάθε νευρώνας του στην συνέχεια καταλήγει σε αυτόν της εξόδου

- **Νευρωνικό Δίκτυο Παλινδρόμησης (General Regression Neural Network)**

- **Πιθανοκρατικό Νευρωνικό δίκτυο (Probabilistic Neural Network – χρησιμοποιείται για οικονομικές εφαρμογές)**

- **Νευρωνικό δίκτυο Ακτινικής Συνάρτησης Βάσης (Radial Basis Function Neural Network)** βασίζεται στην τακτική εύρεσης επιφανείας στον πολυδιάστατο χώρο που είναι ευρέως χρησιμοποιούμενη στην στατιστική και φυσική επιστήμη. Στο περιβάλλον των Τεχνητών Νευρωνικών Δικτύων, οι κρυμμένοι νευρώνες (hidden neurons) του δικτύου παρέχουν ένα σύνολο «συναρτήσεων» που απαρτίζουν μία αυθαίρετη «βάση» για τα

διανύσματα εισόδου και οι οποίες (συναρτήσεις) εξελίσσονται πάνω στα κρυμμένα στρώματα (hidden layers) του δικτύου. Οι συναρτήσεις αυτές καλούνται Ακτινικές Συναρτήσεις Βάσης (Radial-basis functions) και ονοματίζουν αντίστοιχα και το συγκεκριμένο είδος των Τεχνητών Νευρωνικών Δικτύων. Οι Ακτινικές Συναρτήσεις Βάσης πρωτοεμφανίστηκαν για να λύσουν το πρόβλημα παρεμβολής των πολλών μεταβλητών και αποτελούν ένα από τα πιο βασικά μαθηματικά εργαλεία στην επιστήμη της Αριθμητικής Ανάλυσης. Επαναληπτικό Νευρωνικό Δίκτυο (Recurrent Neural Network)

- **Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines)**

Η Μηχανική Μάθηση θεωρείται ως ένα παρακλάδι της Τεχνητής Νοημοσύνης και ασχολείται με

την ανάπτυξη τεχνικών και μεθόδων μέσω του υπολογιστή. Η ανάπτυξη των αλγορίθμων επιτρέπει στον υπολογιστή να ασκεί καθήκοντα και δραστηριότητες.

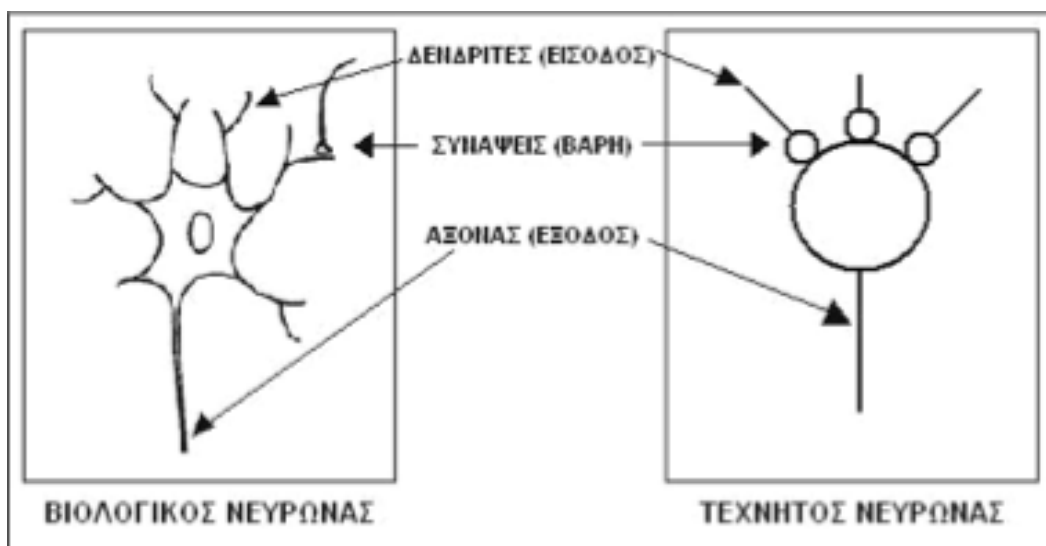
Η Support Vector Machine (SVM) αναπτύχθηκε για πρώτη φορά το 1992, και θεσπίστηκε από τους Boser, Guyon, και Vapnik σε COLT-92. Οι Μηχανές διανύσματος υποστήριξης (SVMs) είναι ένα σύνολο μεθόδων που χρησιμοποιούνται για την ταξινόμηση και παλινδρόμηση.

Ανήκουν σε μια οικογένεια γενικευμένων γραμμικών ταξινομητών. Η Support Vector Machine (SVM) είναι ένα εργαλείο πρόβλεψης, ταξινόμησης και παλινδρόμησης που χρησιμοποιεί τη θεωρία της μηχανικής μάθησης για τη μεγιστοποίηση της προγνωστικής ακρίβειας ενώ αυτόματα

αποφεύγονται τα υπερβολικά ταιριαστά δεδομένα. Η SVM ήταν αρχικά δημοφιλής στην κοινότητα του ερευνητικού κοινού και τώρα είναι ένα ενεργό μέρος της έρευνας της μηχανικής

μάθησης σε όλο τον κόσμο. Η SVM έγινε ευρέως γνωστή όταν έγινε εισαγωγή χρήσης pixel, διότι δίνει ακρίβεια συγκρίσιμη με τα πολύπλοκα νευρωνικά δίκτυα με επεξεργασμένα στοιχεία σε μια εργασία αναγνώρισης χειρογράφου. Έχει, επίσης, χρησιμοποιηθεί από πολλές εφαρμογές, όπως η ανάλυση με το χέρι, η ανάλυση προσώπου και ούτω καθ' εξής, ειδικά για την ταξινόμηση μοτίβου και παλινδρόμησης. Τα θεμέλια της SVM έχουν αναπτυχθεί από τον Vapnik και η SVM χρησιμοποιείται πλέον ευρέως λόγω των πολλών ελπιδοφόρων χαρακτηριστικών, όπως η καλύτερη εμπειρική απόδοση. Η ελαχιστοποίηση του κινδύνου στις SVMs (SRM) έχει αποδειχθεί ότι είναι ανώτερη από την παραδοσιακή εμπειρική ελαχιστοποίηση του κινδύνου (ERM), η οποία χρησιμοποιείται από τα συμβατικά νευρωνικά δίκτυα. Αυτή η διαφορά είναι πολύ σημαντική και εξοπλίζει την SVM με μια μεγαλύτερη δυνατότητα να γενικευθεί. Οι SVMs αναπτύχθηκαν για να λύσουν προβλήματα ταξινόμησης, αλλά πρόσφατα έχουν επεκταθεί στην επίλυση προβλημάτων παλινδρόμησης.

Μετάδοση σήματος μέσα στο νευρωνικό δίκτυο

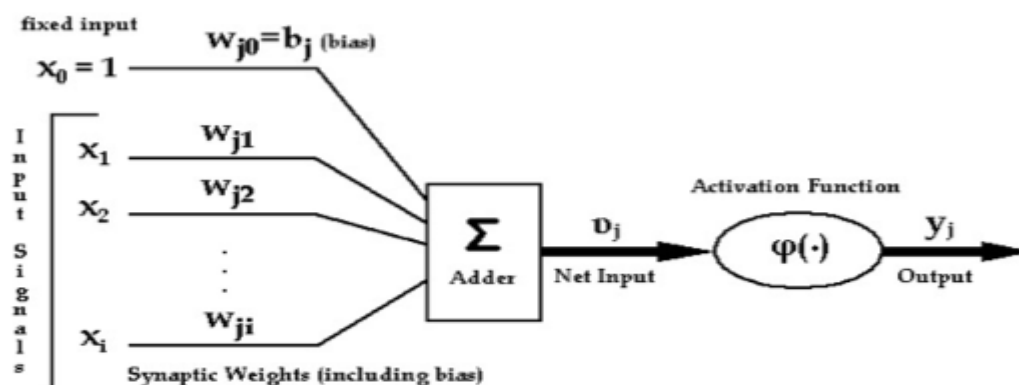


Έχοντας την δομή ενός απλού δικτύου, μένει τώρα να δούμε πως και με ποιά διαδικασία μεταδίδεται το σήμα από νευρώνα σε νευρώνα. Όλα τα σήματα που φθάνουν σε ένα νευρώνα αθροίζονται, υπόκεινται δηλαδή σε μία διαδικασία, παράγεται ως αποτέλεσμα της διαδικασίας μία έξοδος και αυτό είναι το σήμα το οποίο μεταδίδεται περαιτέρω στους επόμενους νευρώνες. Η θεώρηση αυτή είναι γενική και ισχύει πάντοτε, αυτό όμως που αλλάζει είναι η διαδικασία η οποία δεν είναι πάντα η ίδια. Ας δούμε αρχικά δύο τρόπους με τους οποίους γίνεται η μετάδοση αυτή. Ο πρώτος τρόπος είναι δυαδικός. Στην περίπτωση αυτή ένας νευρώνας μπορεί να βρεθεί σε μία από δύο δυνατές καταστάσεις: να είναι ενεργός ή να είναι αδρανής. Όταν ένας νευρώνας δέχεται διάφορα σήματα, τότε υπολογίζει μία ποσότητα x από όλα τα δεδομένα που έχει και συγκρίνει την τιμή της ποσότητας αυτής με μια τιμή κατωφλίου, θ , η οποία είναι χαρακτηριστική (σταθερή)

και ορισμένη από την αρχή για τον νευρώνα αυτόν. Αν η τιμή της ποσότητας είναι μεγαλύτερη από την τιμή κατωφλίου, τότε λέμε ότι ο νευρώνας ενεργοποιείται. Αν όμως είναι μικρότερη, τότε ο νευρώνας παραμένει αδρανής, δηλ. στην δεδομένη στιγμή δεν μεταδίδει κανένα σήμα παραπέρα στο δίκτυο. Επειδή ο νευρώνας εδώ δρα ως δυαδικό στοιχείο, γι' αυτό η έξοδός του, $f(x)$, θα είναι 1 όταν είναι ενεργοποιημένος και 0 όταν είναι αδρανής.

Στο νευρώνα αυτό, η πληροφορία ρέει πάντα προς μία κατεύθυνση, από αριστερά προς τα δεξιά, δεν υπάρχει δηλαδή κανένας βρόχος ανάδρασης. Βάσει αυτού, μπορούμε να διακρίνουμε τρεις βασικές φάσεις της λειτουργίας του

Κατά την πρώτη φάση, κάθε είσοδος πολλαπλασιάζεται με το συναπτικό βάρος που της αντιστοιχεί. Στη δεύτερη φάση οι σταθμισμένες πλέον είσοδοι και ένας εξωτερικά εφαρμοζόμενος παράγοντας, η μεροληψία ή πόλωση ή κατώφλι (*bias, threshold*), αθροίζονται και δίνουν το τοπικό πεδίο (*net input, induced local field, activation potential*). Για λόγους απλούστευσης, η μεροληψία μπορεί να θεωρηθεί ως μία επιπλέον είσοδος, με συναπτικό βάρος ίσο προς την τιμή του και πάγια τιμή εισόδου ίση προς τη μονάδα.



Σχήμα : αναπαράσταση μη γραμμικού νευρώνα

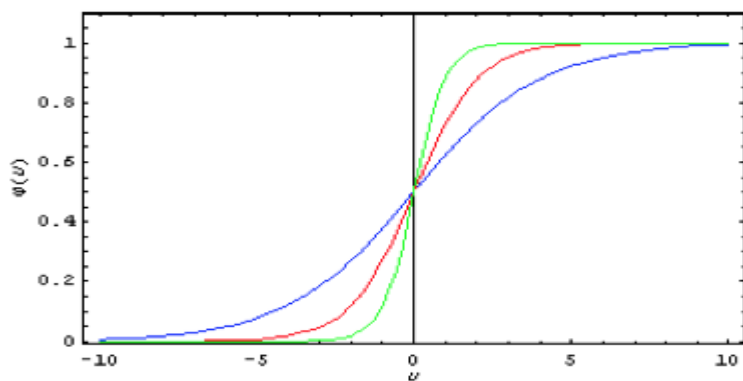
Ως εδώ, ο νευρώνας δεν κάνει τίποτα άλλο από το να δίνει έναν γραμμικό συνδυασμό των εισόδων, με συντελεστές τα προσαρμοζόμενα συναπτικά βάρη. Αν η λειτουργία του λοιπόν σταματούσε εδώ, τότε θα είχαμε έναν γραμμικό νευρώνα, που θα έδινε ένα γραμμικό προσαρμοζόμενο φίλτρο (*linear adaptive filter*). Ένα τεχνητό νευρωνικό δίκτυο που αποτελείται από τέτοιους νευρώνες θα είναι γραμμικό.

Τέλος, στην τρίτη φάση, εφαρμόζεται η *συνάρτηση ενεργοποίησης* ή *συνάρτηση μεταφοράς* (*activation function* ή *squashing function*) στο τοπικό πεδίο και το αποτέλεσμα δίνει την έξοδο του νευρώνα.

Η μετάδοση του σήματος γίνεται πάλι με την συνάρτηση $f(x)$, η οποία τώρα έχει μία ειδική μορφή. Χρησιμοποιούμε όλες τις τιμές των εισόδων και τις τιμές των βαρών, w , και υπολογίζουμε αριθμητικά την $f(x)$. Ένα παράδειγμα μορφής της συνάρτησης αυτής είναι το εξής:

$$f(x) = \frac{1}{1 + e^{-ax}}$$

Η συνάρτηση λέγεται **σιγμοειδής** συνάρτηση. Η γενική της όμως ονομασία σε όλες τις περιπτώσεις είναι συνάρτηση μεταφοράς (*transfer function*), ή συνάρτηση ενεργοποίησης (*activation function*). Το κοινό χαρακτηριστικό που έχουν οι συναρτήσεις αυτές είναι ότι πρέπει να είναι πάντοτε μη-γραμμικές. Δεν αρκούν γραμμικές συναρτήσεις, γιατί τότε η έξοδος θα ήταν ευθέως ανάλογη με την είσοδο, κάτι που δεν μπορεί να συμβεί στα νευρωνικά δίκτυα.

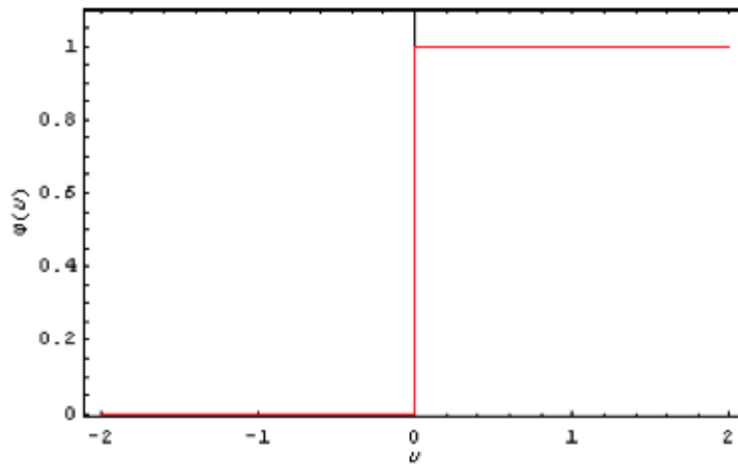


Σχήμα :Η λογιστική συνάρτηση , για $\alpha = 0.5$ (μπλέ) , $\alpha=1$ (κόκκινο) και $\alpha=2$ (πράσινο)

Άλλα παραδείγματα σιγμοειδών συναρτήσεων που χρησιμοποιούνται ως συναρτήσεις ενεργοποίησης είναι η υπερβολική συνάρτηση (hyperbolic function) και η συνάρτηση τόξου εφαπτομένης (arctangent function).

Η **βηματική (step) ή κατωφλίου συνάρτηση** , η οποία δίνει στην έξοδο αποτέλεσμα (συνήθως 1) μόνο αν η τιμή που υπολογίζει ο αθροιστής είναι μεγαλύτερη από μία τιμή κατωφλίου T και εκφράζεται από τη σχέση:

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



Σχήμα :Βηματική συνάρτηση για τιμή κατωφλίου ίση προς μηδέν

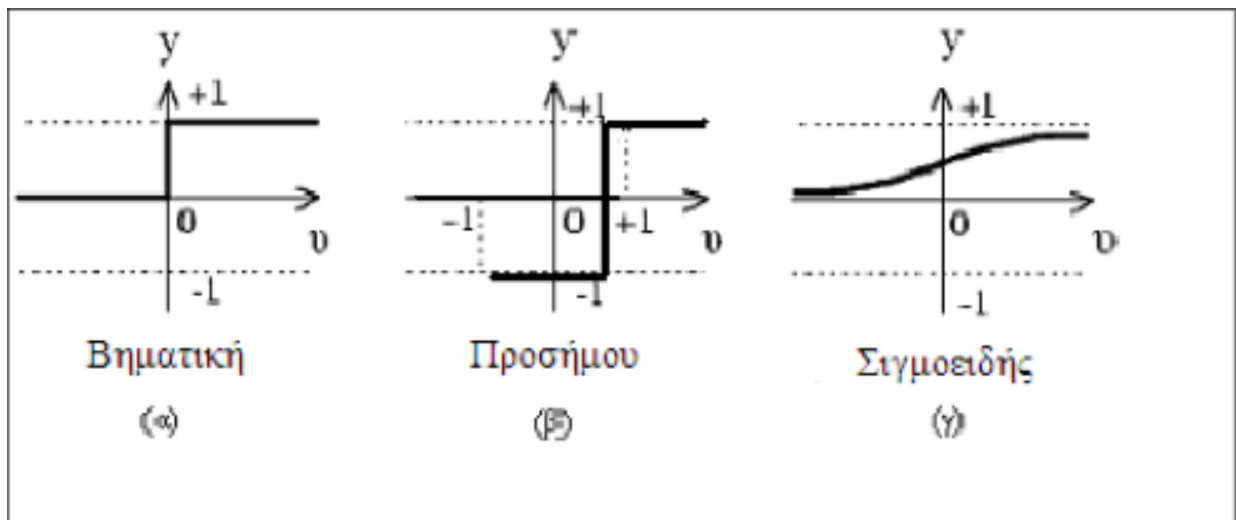
Η συνάρτηση *πρόσημου* (*sign*) η οποία δίνει στην έξοδο αρνητική (ή θετική) πληροφορία αν η τιμή που υπολογίζει ο αθροιστής είναι μικρότερη (ή μεγαλύτερη) από μία τιμή κατωφλίου T . Εκφράζεται από τη σχέση:

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

Νευρώνας υπερβολικής εφαπτομένης (hyperbolic tangent) με την σχέση :

$$F(\sigma) = 1 - e^{-\sigma} / 1 + e^{-\sigma} \in [-1, 1]$$

Perceptron με την σχέση : $f(\sigma) = \text{sgn}(\sigma)$



Σχήμα : Αντιστοιχία φυσικού- τεχνητού νευρώνα

Οι τεχνητοί νευρώνες δίνουν τη δυνατότητα υλοποίησης απλών αλγεβρικών συναρτήσεων, όπως η υλοποίηση με τεχνητό νευρώνα των λογικών συναρτήσεων AND, OR, και NOT. Για παράδειγμα, στην υλοποίηση του NOT χρησιμοποιείται σαν συνάρτηση ενεργοποίησης η βηματική συνάρτηση με κατώφλι $T=-0.5$. Οι τιμές εισόδου μπορούν να κυμαίνονται από 0 (ψευδές) έως 1 (αληθές). Αν η είσοδος του νευρώνα είναι 0 τότε πολλαπλασιαζόμενη με το βάρος $w=-1$ δίνει $S=0$. Η τιμή αυτή ξεπερνά το κατώφλι του -0.5 οπότε στην έξοδο παράγεται 1. Στην περίπτωση που η τιμή εισόδου είναι 1 τότε $S=-1$, τιμή που βρίσκεται κάτω του κατωφλίου -0.5 , με αποτέλεσμα να παράγεται στην έξοδο 0.

ΠΛΕΟΝΕΚΤΗΜΑΤΑ ΚΑΙ ΜΕΙΟΝΕΚΤΗΜΑΤΑ ΤΝΔ

Σε γενικές γραμμές τα ΤΝΔ είναι κατάλληλα να αντιμετωπίσουν συγκεκριμένες κατηγορίες προβλημάτων και ειδικότερα προβλήματα εκπαίδευσης που αντιμετωπίζονται με συσχέτιση προτύπων. Η αντίληψη ότι τα ΤΝΔ μπορούν να λύσουν όλα τα προβλήματα αυτόματης απόκτησης γνώσης είναι μη ρεαλιστική. Η χρήση των Τεχνητών Νευρωνικών Δικτύων παρέχει τα εξής πλεονεκτήματα:

i. Μη γραμμικότητα

Η μη γραμμικότητα οφείλεται στο γεγονός ότι ένα Τεχνητό Νευρωνικό Δίκτυο δομείται από τη σύνδεση νευρώνων, οι οποίοι είναι μη γραμμικές συσκευές. Η μη γραμμικότητα είναι μια πολύ σημαντική ιδιότητα, ιδιαίτερα αν ο φυσικός μηχανισμός για την παραγωγή των σημάτων εισόδου είναι μη γραμμικός.

ii. Σχεδιασμός Εισόδου-Εξόδου

Ο σχεδιασμός εισόδου-εξόδου είναι ένα συνηθισμένο παράδειγμα μάθησης που καλείται επιβλεπόμενη μάθηση, εμπλέκει μεταβολή των βαρών του Νευρωνικού Δικτύου, εφαρμόζοντας ένα σύνολο δειγμάτων εξάσκησης ή παραδείγματα έργων. Κάθε παράδειγμα αποτελείται από ένα μοναδικό σήμα εισόδου και την επιθυμητή απόκριση. Η εξάσκηση του δικτύου επαναλαμβάνεται για πολλά παραδείγματα, μέχρι το δίκτυο να φτάσει σε μια σταθερή κατάσταση, όπου πλέον δεν γίνονται αλλαγές στα βάρη. Έτσι το δίκτυο μαθαίνει από τα παραδείγματα, κατασκευάζοντας ένα σχεδιασμό εισόδου-εξόδου για το πρόβλημα στο χέρι.

iii. **Προσαρμοστικότητα**

Τα Νευρωνικά Δίκτυα έχουν τη δυνατότητα να προσαρμόζουν τα βάρη τους στις αλλαγές του περιβάλλοντός τους . Μερικές φορές οι προσαρμογές οδηγούν σε μείωση της απόδοσης του συστήματος, γι' αυτό θα πρέπει να είναι επαρκείς δίλημμα σταθερότητας-πλαστικότητας.

iv. **Αποδεικτική Απόκριση**

Ένα Νευρωνικό Δίκτυο σχεδιάζεται για να παρέχει πληροφορίες όχι μόνο για το συγκεκριμένο υπόδειγμα που επιλέγεται αλλά και για την εμπιστοσύνη στην απόφαση που παίρνεται.

v. **Συναφής Πληροφορία**

Η γνώση αναπαριστάνεται από την πολύ δομημένη και ενεργή κατάσταση του Νευρωνικού Δικτύου.

vi. **Αντοχή σε σφάλματα**

Τα ΝΔ έχουν μεγάλη ανοχή σε δομικά σφάλματα. Αυτό σημαίνει ότι η κακή λειτουργία ή η καταστροφή ενός νευρώνα ή κάποιων συνδέσεων δεν είναι ικανή να διαταράξει σημαντικά την λειτουργία τους καθώς, όπως αναφέρθηκε, η πληροφορία που εσωκλείουν δεν είναι εντοπισμένη σε συγκεκριμένο σημείο αλλά διάχυτη σε όλο το δίκτυο.

vii. **Υλοποιησιμότητα σε VLSI**

Η συμπαγής παράλληλη φύση του Νευρωνικού Δικτύου, κάνει δυνατή την υλοποίηση

του σε VLSI τεχνολογία, έτσι ώστε τα νευρωνικά δίκτυα να μπορούν να χρησιμοποιηθούν σε εφαρμογές πραγματικού χρόνου.

viii. **Ομοιομορφία Ανάλυσης και Σχεδιασμού**

Η έννοια είναι ότι ο ίδιος συμβολισμός χρησιμοποιείται σε όλα τα πεδία που περιέχουν εφαρμογή των νευρωνικών δικτύων πράγμα που υποδηλώνεται με διάφορους τρόπους: Οι νευρώνες παριστάνουν ένα συστατικό συνηθισμένο σε όλα τα νευρωνικά δίκτυα. Αυτή η ιδιότητα κάνει δυνατό το διαμοιρασμό θεωριών και αλγορίθμων εκμάθησης σε διαφορετικές εφαρμογές των νευρωνικών δικτύων.

ix. **Αναλογία με Νευροβιολογία**

Ο σχεδιασμός νευρωνικών δικτύων γίνεται σε αναλογία με τον εγκέφαλο. Οι νευροβιολόγοι βλέπουν τα νευρωνικά δίκτυα σαν αντικείμενο έρευνας για την εξήγηση νευροβιολογικών φαινομένων. Ομοίως οι μηχανικοί βλέπουν στη νευροβιολογία για νέες ιδέες για την επίλυση πολύπλοκων προβλημάτων.

Συμπεράσματα

Αρχικά στην παρούσα διπλωματική εργασία μελετήθηκαν διάφορες μέθοδοι εξομάλυνσης, οι μέθοδοι πρόβλεψης συνεχούς ζήτησης, οι στατιστικές μέθοδοι πρόβλεψης, τα μοντέλα παλινδρόμησης και η ανάλυση ARIMA οι οποίες χρησιμοποιούνται για την διαμόρφωση βραχυπρόθεσμων ή μακροπρόθεσμων προβλέψεων των μελλοντικών τιμών μιας χρονοσειράς. Οι μέθοδοι εξομάλυνσης γίνονται σε σύντομο χρονικό διάστημα και με χαμηλό κόστος, ενώ βασίζονται σε απλές υπολογιστικές πράξεις και δεν απαιτούν μεγάλο αριθμό παρατηρήσεων της χρονοσειράς. Παρουσιάστηκε η ανάλυση ARIMA η οποία αποτελεί μία στατιστικά εξειδικευμένη προσέγγιση στην ανάλυση και την κατασκευή υποδείγματος πρόβλεψης με σκοπό την την καλύτερη αναπαράσταση μιας χρονοσειράς. Η μέθοδος αυτή αποσπά ένα μεγάλο αριθμό πληροφοριών από τα ιστορικά δεδομένα της χρονοσειράς και των σφαλμάτων.

Η χρήση μοντέλων εύκαμπτης υπολογιστικής θεωρείται προτιμότερη σε συγκριση με αυτή των μεθόδων παλινδρόμησης καθώς τα νευρωνικά και νεύρο-ασαφή μοντέλα, παράγουν καλύτερα αποτελέσματα. Αν και ο χρόνος εκπαίδευσης εξαρτάται από πολλούς παράγοντας, περιορίζεται σε μερικά δευτερόλεπτα, είναι όμως μεγαλύτερος από των μοντέλων AR και ARMA.

Τα νευρωνικά δίκτυα, όπως και τα συστήματα ασαφούς λογικής βρίσκουν ευρεία χρήση, στις επιστημονικές εφαρμογές των τελευταίων ετών. Αξίζει να σημειωθεί ότι τα συστήματα αυτά δεν παρέχουν οριστικά μοντέλα πρόβλεψης, για τον λόγο αυτό λειτουργούν συμπληρωματικά και όχι ανταγωνιστικά των συμβατικών μεθόδων πρόβλεψης. Το γεγονός ότι σε μερικές περιπτώσεις τα νευρωνικά και νεύρο-ασαφή μοντέλα, αποδίδουν χειρότερα από τα μοντέλα παλινδρόμησης, σε καμία περίπτωση δεν μειώνει την αξία τους. Η ασαφής λογική έχει την ικανότητα να

ενσωματώνει την γνώση των ειδικών με άμεσο και εύκολο τρόπο χρησιμοποιώντας κανόνες και γλωσσικές μεταβλητές.

Ο σκοπός της πρόβλεψη είναι να βοηθήσει τη διεύθυνση να λάβει αποφάσεις, να προγραμματίσει τις απαιτήσεις για πρώτες ύλες, τις απαιτήσεις για marketing και γενικά να εξελιχθεί η επιχείρηση μέσα από προβλεπόμενες και στοχευμένες κινήσεις. Για να επιτευχθεί η καλύτερη δυνατή πρόβλεψη θα πρέπει να επιλεγεί η καταλληλότερη μέθοδος προβλέψεων, αντιλαμβανόμαστε λοιπόν την αναγκαιότητα του ερευνητή να διαθέτει εμπειρία οξυδέρκεια, διορατικότητα για να καταλήξει στην κατάλληλη μέθοδο.

ΒΙΒΛΙΟΓΡΑΦΙΑ

1. Armano G, Marchesi M, Murru A, (2005) A hybrid genetic-neural architecture for stock indexes forecasting, *Information Sciences* (170)1: 3–33
2. Ατσαλάκης Γ., Βαλαβάνης Κ., Ζοπουνίδης Κ. (2008), Τεχνικές Χρηματιστηριακών Προβλέψεων, Αθήνα: εκδόσεις κλειδάριθμος
3. Atsalakis George, Valavanis Kimon, (2009). Surveying stock market forecasting techniques – Part II: Soft computing methods. *Expert Systems with Applications* (36) 5932–5941
4. Atsalakis G. & Ucenic C. (2007). Forecasting the tourism demand using ANFIS for assuring successful strategies in the view of sustainable development in the tourism sector. *Proceedings of the 2nd IASME/WSEAS International Conference on Energy & Environment*, (pp. 213-219). Portorose, Slovenia.
5. Atsalakis G. and Valavanis K. (2010). Forecasting stock trends using a combined technical analysis and neuro-fuzzy based approach. *Journal of Financial Decision Making*, 6(1):79-94.
6. Coates Vary, Farooque Mahmud, Klavans Richard, Lapid Koty, Linstone Harold, Pistorius Carl, Porter Alan (2001), On the Future of Technological Forecasting , *Technological Forecasting and Social Change* (67), 1–17
7. Committee on Forecasting Future Disruptive Technologies; National Research Council (2009)
8. Persistent Forecasting of Disruptive Technologies, ISBN: 0-309-11661-9, 136 pages
9. De Gooijer Jan G, Hyndman J. Rob (2006), 25 years of time series forecasting, *Journal of Forecasting*
10. Firat Ayse Kaya, Lee Wei, Woon Stuart Madnick (2008), Technological Forecasting – A Review. Working Paper CISL# 2008-15, Composite Information Systems Laboratory (CISL) Sloan School of Management, Room E53-320 Massachusetts Institute of Technology Cambridge
11. Green Kesten, J. Scott Armstrong (2007), Structured analogies for forecasting Business and Economic Forecasting, *International Journal of Forecasting* (23) 365–376
12. Hasson F, Keeney S, McKenna H (2000). Research guidelines for the Delphi survey technique. *J Adv Nurs*, (32):1008–1015

13. Higgins L. Matthew (2011). *Advances in Economic Forecasting*. Western Michigan University, Upjohn Institute for Employment Research.
14. Hyndman R, Athanasopoulos G, (2013), *Forecasting: Principles and Practice*, <https://www.otexts.org/book/fpp>
15. Κουγιουμτζής Δημήτριος (2005), *Γραμμική ανάλυση χρονοσειρών*, πανεπιστημιακές σημειώσεις
16. Kourentzes Nikolaos, Petropoulos Fotios, Trapero Juan R (2013, June). Improving forecasting by estimating time series structural components across multiple frequencies. Paper presented at ISF, Seoul Korea.
17. Kucharavy Dmitry, Roland De Guio, IAMOT (2008), Technological Forecasting and Assessment of barriers for emerging technologies, International Association for Management of Technology, Proceedings , LGECO - Design Engineering Laboratory, INSA Strasbourg
18. Lawrence Michael, Goodwin Paul, O'Connor Marcus, O'Neil Dilek (2006). Judgmental forecasting: A review of progress over the last 25years. *International Journal of Forecasting*, (22) 493 – 518
19. Μάργια Γεωργία (2009), *Ανάλυση και Πρόβλεψη Χρονοσειρών*, διπλωματική εργασία
20. Μπελλάλη Θ , Καραμήτρη Ι. (2011), Η ερευνητική μεθοδολογία Delphi και η εφαρμογή της στις επιστήμες υγείας, *Archives Of Hellenic Medicine*, 28(6):839-848
21. Melek Acar, Boyacioglu Derya Avci (2010). An Adaptive Network-Based Fuzzy Inference System (ANFIS) for the prediction of stock market return: The case of the Istanbul Stock Exchange *Expert Systems with Applications* (37)7908–7912
22. Moreno-Casbas T, Martín-Arribas C, Orts-Cortés I, Comet- Cortés P. (2011). Investén-Isclii Co-Ordination and Development of Nursing Research Centre. Identification of priorities for nursing research in Spain: A Delphi study. *J Adv Nurs*, 35:857–863
23. Murphy Mk, Black Na, Lamping Dl, Mckee C, Sanderson C, Askham J. (1998) Consensus development methods and their use in clinical guideline development. *Health Technol Asses*, (2):1–88
24. Νέζης Δημήτριος (2010), *Πρόβλεψη τιμών μετοχών με τη χρήση νευρωνικών δικτύων και δεικτών που μετρούν την ψυχολογία των επενδυτών*, μεταπτυχιακή διατριβή
25. Πετρόπουλος Φ., Ασημακόπουλος Β. (2011), *Επιχειρησιακές Προβλέψεις*, Αθήνα: Εκδόσεις Συμμετρία.
26. Reinsel Gregory, Box George, Gwilym M. Jenkins (1977), *Time Series Analysis- Forecasting and Control*. Published by John Wiley & Sons

27. Ρεφανίδης Γιάννης (2011), Νευρωνικά Δίκτυα, Πανεπιστημιακές παραδόσεις για Νευρωνικά Δίκτυα και Εξελικτικούς Αλγορίθμους
28. Σούκος Βασίλειος, Κάντζιος Δημήτριος (2006), Καταγραφή και παρουσίαση παραδοσιακών και σύγχρονων μεθόδων πρόβλεψης , επισκόπηση και εκτίμηση λογισμικών προβλέψεων, διπλωματική εργασία
29. Stone Fish L, Bussy Dm. (2005) The Delphi method. In: Sprenkle DH, Piercy FP (eds). *Research methods in family therapy*, 238–253
30. Yousouf Mi (2007).Using experts’opinions through Delphi technique. *Practical Assessment, Research & Evaluation*, (12):1–8
31. Watson R., Mckenna H, Cowman S, Keady J (2008). Nursing research: Designs and methods. *Churchill Livingstone*, (19) 231–241.
32. Winston W.L, Ibright, S.C. and Zappe, C. (1999) *Data Analysis & Decision Making with Microsoft Excel*, Duxbury Press, USA.
33. Zhang Guoqiang, Patuwo Eddy, Michael Y. Hu (1997) Forecasting with artificial neural networks: The state of the art, *International Journal of Forecasting* (14)1: 1-156
34. Zhang G. Peter (2003), Time series forecasting using a hybrid ARIMA and neural network model, *Neurocomputing* (50) 159 – 175