



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ
ΕΠΙΣΤΗΜΩΝ

ΔΙΑΜΕΤΑΠΤΥΧΙΑΚΟ - ΔΙΑΤΜΗΜΑΤΙΚΟ
ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ Δ.Π.Μ.Σ.
“ΜΑΘΗΜΑΤΙΚΗ ΠΡΟΤΥΠΟΠΟΙΗΣΗ ΣΤΙΣ ΣΥΓΧΡΟΝΕΣ ΤΕΧΝΟΛΟΓΙΕΣ
ΚΑΙ ΤΗΝ ΟΙΚΟΝΟΜΙΑ”

ΑΞΙΟΛΟΓΗΣΗ ΑΠΟΔΟΤΙΚΟΤΗΤΑΣ ΕΙΣΗΓΜΕΝΩΝ ΕΤΑΙΡΕΙΩΝ
ΜΕ ΕΥΦΥΕΙΣ ΥΠΟΛΟΓΙΣΤΙΚΕΣ ΜΕΘΟΔΟΥΣ

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΑΤΡΙΒΗ
Δρ. ΔΗΜΗΤΡΙΟΥ Χ. ΜΑΝΤΖΑΡΗ
ΔΙΠΛΩΜΑΤΟΥΧΟΥ ΗΛΕΚΤΡΟΛΟΓΟΥ ΜΗΧΑΝΙΚΟΥ

Επιβλέπων Καθηγητής: Δρ Ιωάννης Ντόκας

ΑΘΗΝΑ, ΑΠΡΙΛΙΟΣ 2013



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ
ΕΠΙΣΤΗΜΩΝ

ΔΙΑΜΕΤΑΠΤΥΧΙΑΚΟ - ΔΙΑΤΜΗΜΑΤΙΚΟ
ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ Δ.Π.Μ.Σ.
“ΜΑΘΗΜΑΤΙΚΗ ΠΡΟΤΥΠΟΠΟΙΗΣΗ ΣΤΙΣ ΣΥΓΧΡΟΝΕΣ ΤΕΧΝΟΛΟΓΙΕΣ
ΚΑΙ ΤΗΝ ΟΙΚΟΝΟΜΙΑ”

ΑΞΙΟΛΟΓΗΣΗ ΑΠΟΔΟΤΙΚΟΤΗΤΑΣ ΕΙΣΗΓΜΕΝΩΝ ΕΤΑΙΡΕΙΩΝ
ΜΕ ΕΥΦΥΕΙΣ ΥΠΟΛΟΓΙΣΤΙΚΕΣ ΜΕΘΟΔΟΥΣ

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΑΤΡΙΒΗ
Δρ. ΔΗΜΗΤΡΙΟΥ Χ. ΜΑΝΤΖΑΡΗ
ΔΙΠΛΩΜΑΤΟΥΧΟΥ ΗΛΕΚΤΡΟΛΟΓΟΥ ΜΗΧΑΝΙΚΟΥ

Επιβλέπων Καθηγητής: Δρ Ιωάννης Ντόκας

ΑΘΗΝΑ, ΑΠΡΙΛΙΟΣ 2013

ΕΥΧΑΡΙΣΤΙΕΣ

Από τη θέση αυτή θα ήθελα να ευχαριστήσω όλους όσους συνέβαλλαν στην ολοκλήρωση της διατριβής αυτής και ιδιαίτερα:

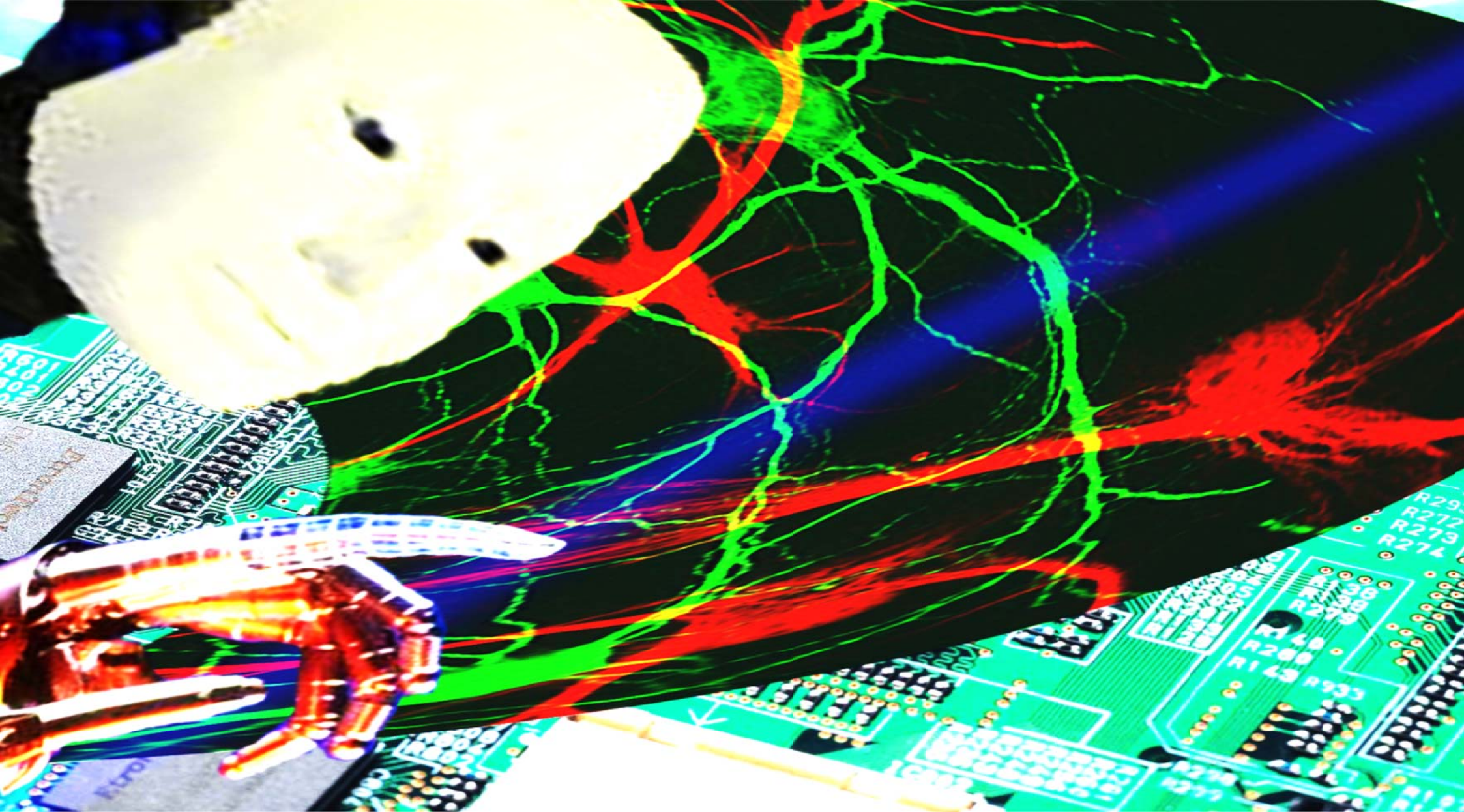
Θερμά ευχαριστώ τον επιβλέποντα της μεταπτυχιακής διατριβής, Δρ. Ιωάννη Ντόκα, Λέκτορα (ΠΔ 407/80) του Τμήματος Οικονομικών Επιστημών, Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών και της Σχολής Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών, Εθνικού Μετσόβιου Πολυτεχνείου, πρόεδρο της τριμελούς επιτροπής, για τη δυνατότητα εκπόνησης της συγκεκριμένης διατριβής, την καθοδήγηση σε κρίσιμες περιόδους, στα ευχάριστα, αλλά και στα δύσκολα, τη διάθεση σημαντικού μέρους του χρόνου του, καθώς και για τη γενικότερη συνεργασία μας όλα αυτά τα χρόνια.

Τον Δρ. Απόστολο Χριστόπουλο, Λέκτορα (ΠΔ 407/80) του Τμήματος Οικονομικών Επιστημών, Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών για την άριστη συνεργασία μας, τη συμβολή του και τις εποικοδομητικές συμβουλές του καθ' όλη τη διάρκεια του ερευνητικού έργου.

Τον κ. Λεβεντίδη Ιωάννη, Επίκουρο Καθηγητή του Τμήματος Οικονομικών Επιστημών, Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών για την τιμή που μου έκανε να συμμετάσχει ως μέλος της τριμελούς εξεταστικής επιτροπής.

Καθοριστικό ρόλο για την εκπόνηση της παρούσας μεταπτυχιακής διατριβής έπαιξαν οι γονείς μου, Χαράλαμπος και Αναστασία, οι οποίοι με στήριξαν σε όλη την πορεία μου.

Κλείνοντας, θέλω να ευχαριστήσω όσους ήταν δίπλα μου σε αυτό το ταξίδι της γνώσης.



Abstract

The financial engineering is a dynamic and rapidly evolving research area. The new technologies developed by engineers, computer scientists and mathematicians are widely known and applied in various fields to solve real world problems.

The financial analysis is field of the Financial and focuses on the study of the relationships that exist between the economic balance sheet and the income statement (financial statements) of a company at a given time. This analysis of accounting statements aims to inferences, based on scientific criteria, which will be driven in making economic decisions. The financial analysis crowd interested individuals and financial institutions, and therefore is a valuable tool.

Nowadays, there are specific efforts in order to develop modern tools, based on the available counting data, that will be sufficient in financial evaluation of firms. In the present study, it was attempted the implementation of intelligent computational methods based on artificial neural networks to assess the profitability of listed firms. The implemented algorithms of this thesis are the first attempts for intelligent methods in this financial field.

Various architectures of artificial neural networks implemented and evaluated for proper assessment of the effectiveness of the aforementioned firms, using ratios as well as net values of the financial statements. The architecture with the best results is the proposed artificial neural network and can be used by experts as a parallel auxiliary tool.

The obtained results proofs the ability of artificial neural networks in order to be used in the field of financial analysis. This fact does not abolish the human role by any machine. Simply, the continuous technological development will greatly facilitate the work of the special and contribute to faster against various undesirable for the company statements. The human brain will have the master role, as it would not be ignored that the machines and computers are human contrivances.

ΕΙΣΑΓΩΓΗ

Η εφαρμογή μεθόδων από την επιστήμη του μηχανικού στη χρηματοοικονομική (financial engineering) είναι ένας δυναμικός και ταχύτατα εξελισσόμενος τομέας έρευνας. Οι νέες τεχνολογίες που αναπτύσσονται από τους μηχανικούς, τους επιστήμονες της πληροφορικής και τους μαθηματικούς, γίνονται ευρέως γνωστές και εφαρμόζονται σε διάφορους τομείς για την επίλυση προβλημάτων που αφορούν όλο το εύρος των επιστημών.

Αναφορικά με τα ζητήματα που καλείται να αντιμετωπίσει η οικονομική επιστήμη, η μαθηματική διατύπωση του κάθε προβλήματος αποτελεί την αφετηρία για την επίτευξη επιστημονικά αποδεκτής λύσης. Η οικονομική επιστήμη αποτελείται από διάφορους επιστημονικούς κλάδους ανάλογα με την φύση και το είδος του προβλήματος που επιχειρεί να επιλύσει. Ένας από τους σημαντικότερους κλάδους αποτελεί αυτός της Χρηματοοικονομικής και Λογιστικής, όπου βασικό πεδίο μελέτης αποτελεί η βιωσιμότητα των παραγωγικών οικονομικών μονάδων. Ειδικότερα, το ερευνητικό ενδιαφέρον εστιάζεται στις οικονομικές μονάδες που έχουν σαν βασικό στόχο ύπαρξης και λειτουργίας την μεγιστοποίηση του κέρδους.

Η Λογιστική σαν κλάδος της οικονομικής επιστήμης έχει την ευθύνη παροχής οικονομικών, κατά κανόνα, πληροφοριών οι οποίες αφορούν την λειτουργία των επιχειρήσεων και κοινοποιούνται προς κάθε ενδιαφερόμενο μέσα από ειδικούς πληροφοριακούς πίνακες που καλούνται λογιστικές καταστάσεις (Financial Statements). Η περαιτέρω επεξεργασία της λογιστικής πληροφόρησης αποτελεί μια ουσιαστική διαδικασία για τη λήψη αποφάσεων η οποία καλείται χρηματοοικονομική ανάλυση. Ως το πλέον δημοφιλές εργαλείο χρηματοοικονομικής ανάλυσης έχει αναδειχθεί η ανάλυση αριθμοδεικτών (Ratio Analysis), η οποία παρά τους περιορισμούς που συνοδεύουν τη χρήση της, δύναται να προσφέρει στον αναλυτή αντικειμενικές εκτιμήσεις υπό την προϋπόθεση ότι υπάρχει αντίληψη αυτών των περιορισμών και λαμβάνονται υπόψη στην ακολουθούμενη προσέγγιση ανάλυσης και ερμηνείας των αποτελεσμάτων. Παρόλα αυτά η επιθυμία για βελτίωση της αντικειμενικότητας των εκτιμήσεων που προσφέρουν οι δείκτες, οδήγησαν την επιστημονική έρευνα στην χρήση ποσοτικών μεθόδων και κατ' επέκταση στη δημιουργία υποδειγμάτων εκτίμησης ή και πρόβλεψης μιας κατάστασης η οποία αφορά τη θέση μιας επιχείρησης.

Σήμερα γίνονται ιδιαίτερες προσπάθειες ανάπτυξης σύγχρονων εργαλείων, που βάση των διαθέσιμων στοιχείων των λογιστικών καταστάσεων, θα είναι επαρκή στην χρηματοοικονομική αξιολόγηση των επιχειρήσεων. Στα πλαίσια της παρούσας μεταπτυχιακής διατριβής επιχειρήθηκε η ανάπτυξη ευφών υπολογιστικών μεθόδων που βασίζονται στα τεχνητά νευρωνικά δίκτυα για την αποτίμηση της αποδοτικότητας των εισηγμένων στο χρηματιστήριο εταιρειών. Οι αλγόριθμοι που υλοποιούνται στη διατριβή αυτή, αποτελούν τις πρώτες προσπάθειες εφαρμογής ευφών μεθόδων στο συγκεκριμένο χρηματοοικονομικό τομέα.

Ποικίλες αρχιτεκτονικές τεχνητών νευρωνικών δικτύων υλοποιούνται και αξιολογούνται για την ορθή εκτίμηση της αποτελεσματικότητας των προαναφερθέντων εταιρειών. Η αρχιτεκτονική με τα καλύτερα αποτελέσματα είναι η επικρατούσα και μπορεί να χρησιμοποιηθεί από τους ειδικούς σαν παράλληλο βοηθητικό μέσο.

Τα αποτελέσματα που προκύπτουν, αποδεικνύουν την ικανότητα των τεχνητών νευρωνικών δικτύων να χρησιμοποιηθούν στο πεδίο της χρηματοοικονομικής ανάλυσης. Αυτό δεν σημαίνει ότι καταργείται ο ρόλος του ανθρώπου και αντικαθίσταται από οποιοδήποτε μηχάνημα. Απλώς, η συνεχής τεχνολογική ανάπτυξη θα διευκολύνει τα μέγιστα τη δουλειά του ειδικού και θα συμβάλλει στην ταχύτερη αντιμετώπιση ποικίλων μη επιθυμητών για την εταιρεία καταστάσεων. Πάντοτε κυρίαρχο ρόλο θα κατέχει ο ανθρώπινος νους και αυτός θα δίνει τις κατευθυντήριες γραμμές. Άλλωστε δε θα πρέπει να παραβλέπεται ότι οι μηχανές και οι υπολογιστές αποτελούν ανθρώπινα δημιουργήματα.

Η δομή της μεταπτυχιακής διατριβής εκτείνεται σε πέντε κεφάλαια, ως εξής:

Στο Κεφάλαιο 1 επιχειρείται εισαγωγή στη χρησιμοποίηση μεθόδων από την επιστήμη του μηχανικού στο χρηματοοικονομικό τομέα. Παράλληλα γίνεται μια παρουσίαση της Τεχνητής – Υπολογιστικής Νοημοσύνης και της ανάπτυξης ευφύων συστημάτων.

Στο Κεφάλαιο 2 παρουσιάζονται οι χρηματοοικονομικές μέθοδοι ανάλυσης λογιστικών καταστάσεων των εταιρειών. Ιδιαίτερη έμφαση δίνεται στην ανάλυση με αριθμοδείκτες και αναλύονται οι διαθέσιμοι αριθμοδείκτες.

Στο Κεφάλαιο 3 παρουσιάζονται τα Τεχνητά Νευρωνικά Δίκτυα (ΤΝΔ), οι αρχιτεκτονικές, οι βασικές τους λειτουργίες, οι αλγόριθμοι εκπαίδευσης και οι συναρτήσεις μεταφοράς τους.

Στο Κεφάλαιο 4 περιγράφονται οι μεθοδολογίες των αναπτυχθέντων ΤΝΔ που χρησιμοποιήθηκαν για την αξιολόγηση της αποδοτικότητας των εισηγμένων στο χρηματιστήριο εταιρειών. Τα προτεινόμενα ΤΝΔ στηρίζονται στα πολυ-στρωματικά νευρωνικά δίκτυα (Multi-layer Perceptrons – MLPs) και στα πιθανοκρατικά νευρωνικά δίκτυα (Probabilistic Neural Networks – PNNs). Οι υλοποιήσεις και τα αποτελέσματα που προέκυψαν από την εφαρμογή τους είναι πρωτότυπες και παρουσιάζονται στο κεφάλαιο αυτό.

Στο Κεφάλαιο 5 παρατίθενται γενικά συμπεράσματα και προοπτικές αναφορικά με τις αναπτυχθείσες τεχνικές, και τη δυνατότητα επέκτασής τους σε άλλους τομείς της Χρηματοοικονομικής και των Οικονομικών.

Αθήνα, 16.03.2013
Δημήτριος Χ. Μαντζάρης

ΠΕΡΙΕΧΟΜΕΝΑ

ΚΕΦΑΛΑΙΟ 1^ο

ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗ ΜΗΧΑΝΙΚΗ	11
1.1 Εισαγωγή.....	11
1.2 Επιχειρηματική Δραστηριότητα & Οικονομικό Περιβάλλον.....	11
1.3 Αποδοτικότητα Επιχειρήσεων & Υπολογιστική Νοημοσύνη.....	12
ΒΙΒΛΙΟΓΡΑΦΙΑ ΚΕΦΑΛΑΙΟΥ	14

ΚΕΦΑΛΑΙΟ 2^ο

ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗ ΑΝΑΛΥΣΗ	15
2.1 Εισαγωγή.....	15
2.2 Μέθοδοι Ανάλυσης Λογιστικών Καταστάσεων.....	16
2.2.1 Διαστρωματική ή Κάθετη Ανάλυση.....	16
2.2.2 Συγκριτική ή Διαχρονική Ανάλυση.....	17
2.2.3 Εξειδικευμένες Μέθοδοι – Ανάλυση Νεκρού Σημείου.....	19
2.3 Ανάλυση Λογιστικών Καταστάσεων με Χρήση Αριθμοδεικτών.....	21
2.3.1 Αριθμοδείκτες Ρευστότητας.....	21
2.3.2 Αριθμοδείκτες Δραστηριότητας.....	22
2.3.3 Αριθμοδείκτες Αποδοτικότητας.....	23
2.3.4 Αριθμοδείκτες Δαπανών Λειτουργίας.....	23
2.3.5 Αριθμοδείκτες Διάρθρωσης Κεφαλαίων και Βιωσιμότητας.....	23
2.3.6 Επενδυτικοί Αριθμοδείκτες.....	26
2.4 Περιορισμοί και Αδυναμίες της Ανάλυσης με Χρηματοοικονομικούς Δείκτες.....	26
ΒΙΒΛΙΟΓΡΑΦΙΑ ΚΕΦΑΛΑΙΟΥ	29

ΚΕΦΑΛΑΙΟ 3^ο

ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ	31
3.1 Εισαγωγή.....	31
3.2 Ιστορική Αναδρομή.....	34
3.3. Μοντέλο Τεχνητού Νευρώνα.....	36
3.4 Συναρτήσεις Μεταφοράς.....	38
3.5 Τοπολογίες Νευρωνικών Δικτύων.....	40
3.5.1 Επίπεδα Νευρωνικού Δικτύου.....	40
3.5.2 Στατικά και Δυναμικά Νευρωνικά Δίκτυα.....	41
3.5.3 Αισθητήριο.....	41
3.5.4 Πολλαπλών Στρωμάτων Νευρωνικά Δίκτυα.....	43

3.5.5 Δίκτυα Ακτινικής Βάσης	44
3.5.6 Αυτο-Οργανώμενα Δίκτυα	48
3.5.7 Δίκτυο Elman	51
3.5.8 Δίκτυο Hopfield	52
3.6 Εκπαίδευση Τεχνητού Νευρωνικού Δικτύου	53
3.6.1 Επιβλεπόμενη Μάθηση.....	54
3.6.2 Μη-Επιβλεπόμενη Μάθηση.....	55
3.7 Κανόνες Εκμάθησης Τεχνητού Νευρωνικού Δικτύου	55
3.7.1 Κανόνας του Hebb (Hebb's Rule).....	55
3.7.2 Αλγόριθμος Ελάχιστου Μέσου Τετραγωνικού Σφάλματος.....	56
3.7.3 Αλγόριθμος Πτωτικής Κλίσης.....	57
3.7.4 Αλγόριθμος Οπισθοδιάδοσης	58
3.7.5 Αλγόριθμος Πτωτικής Κλίσης με Ροπή.....	59
3.7.6 Πτωτική Κλίση με Μεταβλητό Ρυθμό Εκμάθησης (Gradient Descent with Variable Learning Rate)	59
3.7.7 Ευπροσάρμοστη Οπισθοδιάδοση	59
3.7.8 Αλγόριθμος Συζευγμένης Κλίσης	60
3.7.9 Αλγόριθμος Quasi-Newton.....	61
3.7.10 Αλγόριθμος One Step Secant	61
3.7.11 Αλγόριθμος Levenberg-Marquardt	62
3.7.12 Κανόνας Εκμάθησης Kohonen.....	63
3.8 Εφαρμογές Νευρωνικών Δικτύων.....	64
ΒΙΒΛΙΟΓΡΑΦΙΑ ΚΕΦΑΛΑΙΟΥ.....	66

ΚΕΦΑΛΑΙΟ 4^ο

ΕΥΦΥΕΙΣ ΥΠΟΛΟΓΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ ΣΤΗΝ ΑΞΙΟΛΟΓΗΣΗ ΤΗΣ ΑΠΟΔΟΤΙΚΟΤΗΤΑΣ ΕΙΣΗΓΜΕΝΩΝ ΕΤΑΙΡΕΙΩΝ 79

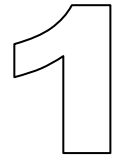
4.1 Εισαγωγή.....	79
4.2 Συλλογή Δεδομένων.....	79
4.3 Νευρωνικά Δίκτυα για την Ευφυή Αξιολόγηση της Εταιρικής Αποδοτικότητας.....	84
4.4 Αποτελέσματα Προσομοίωσης	91
ΒΙΒΛΙΟΓΡΑΦΙΑ ΚΕΦΑΛΑΙΟΥ.....	97

ΚΕΦΑΛΑΙΟ 5^ο

ΣΥΜΠΕΡΑΣΜΑΤΑ ΠΡΟΟΠΤΙΚΕΣ 99

ΓΛΩΣΣΑΡΙ 101

ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗ ΜΗΧΑΝΙΚΗ



1.1 Εισαγωγή

Η έννοια της χρηματοοικονομικής διοίκησης μιας επιχείρησης αναφέρεται στη χρήση των οικονομικών της πόρων, που είναι τα μετρητά και τα περιουσιακά της στοιχεία, με στόχο τη λήψη αποφάσεων που θα καθορίσουν το μέλλον της. Οι αποφάσεις αυτές αφορούν τα χρήματα, που ονομάζονται κεφάλαιο ή κεφάλαια και τα οποία ανέρχονται σε ένα συγκεκριμένο ποσό. Οι ακολουθούμενες επενδυτικές επιλογές, που είναι αποτέλεσμα των ληφθέντων αποφάσεων, απαιτείται να έχουν σαν γνώμονα την εξασφάλιση της βιωσιμότητας της επιχείρησης καθώς και την αύξηση της αξίας της.

Ο ρόλος των μηχανικών είναι καθοριστικής σημασίας στη λήψη επενδυτικών αποφάσεων λόγω της ικανότητας και της εμπειρίας τους στους τομείς της ανάλυσης προβλημάτων και της σχεδίασης και ανάπτυξης αποδοτικών λύσεων. Αυτό είχε σαν αποτέλεσμα την ανάπτυξη του κλάδου της εφαρμογής των μεθόδων των μηχανικών στη χρηματοοικονομική (financial engineering). Ο κλάδος του financial engineering αποτελεί ένα διεπιστημονικό πεδίο που περιλαμβάνει οικονομικές θεωρίες, τεχνικές μεθόδους της επιστήμης του μηχανικού και προγραμματιστικές εφαρμογές [1, 2].

1.2 Επιχειρηματική Δραστηριότητα & Οικονομικό Περιβάλλον

Η παγκόσμια οικονομική ύφεση του 2008 είναι το αποτέλεσμα που προκλήθηκε από τη διεθνή χρηματοπιστωτική κρίση του 2007 σε τομείς της πραγματικής οικονομίας της διεθνούς κοινότητας. Η πιστωτική κρίση από την έλλειψη ρευστότητας έγινε εμφανής στους χρηματιστηριακούς δείκτες των ανεπτυγμένων χωρών, ενώ πολλά χρηματιστήρια παρέμειναν κλειστά λόγω των μεγάλων απωλειών από τις πωλήσεις των επενδυτών [3].

Την λεγόμενη “Μαύρη Παρασκευή”, στις 24 Οκτωβρίου 2008, σημειώθηκαν απώλειες \$3.2 τρις στην Ευρασία, ενώ ξένοι επενδυτές έσπευσαν σε μαζικές ρευστοποιήσεις στην Ελλάδα. Ο Γενικός Δείκτης του Χρηματιστηρίου Αξιών Αθηνών (Χ.Α.Α.) έπεσε κάτω από τις 1.800 μονάδες, στα χαμηλότερα επίπεδα από το 1989, παρά τα πραγματικά κέρδη που εμφάνιζαν οι ελληνικές τράπεζες και επιχειρήσεις. Στην Ελλάδα σημειώθηκαν απώλειες 24 δις. ευρώ στο δεύτερο εξάμηνο του 2008, μέσα σε 24 μέρες [3].

Σημαντικοί κλάδοι της οικονομίας επλήγησαν, όπως ο τραπεζικός, ο ασφαλιστικός, ο κτηματομεσιτικός, οι αυτοκινητοβιομηχανίες, το εμπόριο και άλλοι, με άμεσα αποτελέσματα στα οικονομικά μεγέθη των οικονομιών, όπως η αύξηση της ανεργίας, οι νέες πληθωριστικές πιέσεις, η υποτίμηση νομισμάτων κ.α. Σύμφωνα με αναλυτές, η κατάσταση της οικονομικής ύφεσης προσομοιάζει με την παγκόσμια οικονομική ύφεση του 1929, με τους ρυθμούς ανάπτυξης να υποχωρούν και τα δημοσιονομικά των χωρών να εκτίθενται στο ενδεχόμενο εκτροπής.

Η κρίση έγινε σύντομα εμφανής σε μεγέθη της πραγματικής ελληνικής οικονομίας. Τα έσοδα του κράτους ανήλθαν στα 39.2 δισ. ευρώ, έναντι στόχου 41.6 δισ. ευρώ, ενώ στην Ελλάδα σημειώθηκαν οι μεγαλύτερες απώλειες βιομηχανικών παραγγελιών στην Ευρωζώνη, ύψους 10%, τη στιγμή που εμφανίστηκαν απώλειες 7% στη Γαλλία. Αναλυτές προέβλεψαν μείωση της κατανάλωσης, των εξαγωγών και της παραγωγικής δραστηριότητας, ενώ οι κατασκευές παρέμεναν και παραμένουν παγωμένες. Απαισιόδοξες ήταν οι προβλέψεις του οικονομικού Τύπου και για τα ποσοστά της ανεργίας στον κόσμο [3].

Ο τομέας της ελληνικής επιχειρηματικής δραστηριότητας επηρεάστηκε σημαντικά από την εν λόγω χρηματοπιστωτική κρίση. Συγκεκριμένα, η πτώση της ζήτησης των προϊόντων και υπηρεσιών, η πίεση για χαμηλότερες τιμές, η μεγαλύτερη πίστωση αλλά και τα έντονα προβλήματα είσπραξης απαιτήσεων και εξόφλησης υποχρεώσεων με παράλληλη ελλιπή χρηματοδότηση από τις τράπεζες είναι μια σειρά δυσμενών παραγόντων για τις επιχειρήσεις.

1.3 Αποδοτικότητα Επιχειρήσεων & Υπολογιστική Νοημοσύνη

Η παρουσίαση στα προηγούμενα εδάφια καθιστά σαφές ότι είναι επιτακτική η χρηματοοικονομική αξιολόγηση των επιχειρήσεων με στόχο τη βιωσιμότητα τους και τη διατήρηση ή και την αύξηση της αξίας τους.

Η εν λόγω αξιολόγηση μπορεί να επιτευχθεί με μεθόδους χρηματοοικονομικής ανάλυσης, όπως η κάθετη και η οριζόντια ανάλυση και η ανάλυση με αριθμοδείκτες [4, 5, 6]. Από την άλλη πλευρά, η εμφάνιση των ηλεκτρονικών υπολογιστών και εξέλιξή τους με ταχύτατους ρυθμούς έχουν δώσει ώθηση στις μεθόδους της επιστήμης του μηχανικού για εφαρμογή στη χρηματοοικονομική ανάλυση (financial engineering) [1, 2]. Συγκεκριμένα, η Υπολογιστική Νοημοσύνη (Computational Intelligence – CI) είναι ένα από τα νεότερα επιστημονικά πεδία, που στόχος της είναι, πέρα από την κατανόηση του τρόπου λειτουργίας του ανθρώπινου εγκεφάλου, η κατασκευή νοημόνων οντοτήτων [7-9].

Η Υπολογιστική Νοημοσύνη έχει σαν κύριο επιστημονικό σκοπό της την κατανόηση των φυσικών και των τεχνητών συστημάτων. Η Υπολογιστική Νοημοσύνη συνδυάζοντας στοιχεία μάθησης, προσαρμογής (adaptation), εξέλιξης (evolution) και ασαφούς λογικής (fuzzy logic)

μπορεί να σχεδιάζει συστήματα και προγράμματα, τα οποία εκτελούν διεργασίες που κατά μια έννοια είναι ευφυείς [9-10].

Η ΤΝ διαιρείται στη Συμβολική Τεχνητή Νοημοσύνη, η οποία επιχειρεί να εξομοιώσει την ανθρώπινη νοημοσύνη αλγοριθμικά χρησιμοποιώντας σύμβολα και λογικούς κανόνες, και στην Υπολογιστική Νοημοσύνη, η οποία προσπαθεί να αναπαράγει την ανθρώπινη νοημοσύνη χρησιμοποιώντας μαθηματικές μεθόδους που προσομοιώνουν πραγματικές βιολογικές διαδικασίες (όπως η εξέλιξη των ειδών και η λειτουργία του εγκεφάλου) [11].

Η παρούσα διατριβή επιχειρεί να αναπτύξει ευφυείς υπολογιστικές μεθόδους βασισμένους στα Τεχνητά Νευρωνικά Δίκτυα (ΤΝΔ) που θα αποτελέσουν χρήσιμα εργαλεία στην αξιολόγηση της αποδοτικότητας εισηγμένων εταιρειών στο Χρηματιστήριο Αξιών Αθηνών (Χ.Α.Α). Η αξιοπιστία των προτεινόμενων μοντέλων αξιολογήθηκε χρησιμοποιώντας σαν κριτήρια την ακρίβεια ταξινόμησης (classification accuracy), που εκφράζει το ποσοστό των σωστά ταξινομημένων δειγμάτων σε σχέση με το σύνολο. Για την αξιολόγηση των αναπτυχθέντων μοντέλων, εκτός από τα δεδομένα με τα οποία έγινε η εκπαίδευσή τους, χρησιμοποιήθηκαν και νέα στοιχεία με στόχο την εκτίμηση της συμπεριφοράς τους όταν δεχθούν στις εισόδους τους καινούρια, άγνωστα δεδομένα σε πραγματικές συνθήκες.

ΒΙΒΛΙΟΓΡΑΦΙΑ ΚΕΦΑΛΑΙΟΥ

1. Neftci S., “Principles of Financial Engineering”, 2nd edition, *Academic Press Elsevier*, 2008 (ISBN-13: 978-0-12-373574-4).
2. Blank L., and Tarquin A., “Engineering Economy”, *McGraw-Hill*, 2012 (ISBN 978-0-07-337630-1).
3. Petrakis P., “The Greek Economy and the Crisis, Challenges and Responses”, *Springer-Verlag Berlin Heidelberg*, 2012 (ISBN 978-3-642-21174-4).
4. Vernimmen P., Quiry P., Dallochio M., Le Fur Y., and Salvi A., “Corporate Finance Theory and Practice”, 2nd edition, *John Wiley & Sons*, 2009 (ISBN 978-0-470-72192-6).
5. Atrill P., and McLaney E., “Accounting and Finance for Non-Specialists”, 5th edition, *Prentice Hall*, 2006 (ISBN-13: 978-0-273-70244-3).
6. Dyson J., “Accounting for Non-Accounting Students”, 8th edition, *Financial Times, Prentice Hall*, 2010 (ISBN: 978-0-273-72297-7).
7. Ritter H., Martinetz T., and Schulten K. “Neural Computation and Self-Organizing Maps - An Introduction”, *Addison - Wesley New York*, 1992 (ISBN-10: 0201554429, ISBN-13: 978-0201554427).
8. Charniak E., and McDermott D., “Introduction to Artificial Intelligence” *Addison-Wesley, Reading, MA*, 1985.
9. Winston H., “Artificial Intelligence” *Addison-Wesley*, 2nd ed. *Reading MA*, 1994.
10. Luger F., and Stubblefield A., “Artificial Intelligence: Structures and Strategies for Complex Problem Solving”, *Benjamin/Cummings, Menlo Park CA*, 1993.
11. Βλαχάβας Ι., Κεφαλάς Π., Βασιλειάδης Ν., Κόκκορας Φ., και Σακελλαρίου Η., “Τεχνητή Νοημοσύνη”, Γ' Έκδοση, *Εκδόσεις Β. Γκιούρδας Εκδοτική*, 2006 (ISBN: 960-387-431-0).

ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗ ΑΝΑΛΥΣΗ

2

2.1 Εισαγωγή

Ένας από τους βασικότερους σκοπούς της λογιστικής είναι η σύνταξη των λογιστικών ή χρηματοοικονομικών καταστάσεων των επιχειρήσεων και οργανισμών. Εντούτοις, η διερεύνηση, η ερμηνεία και η αξιολόγηση των στοιχείων των προαναφερθέντων καταστάσεων είναι πολύ σημαντικές ενέργειες καθώς μπορούν να παρέχουν ωφέλιμες πληροφορίες προς τους ενδιαφερόμενους των επιχειρηματικών μονάδων ώστε να λάβουν τις ορθές αποφάσεις. Παρόλο που οι χρηματοοικονομικές καταστάσεις αποτελούν σημαντική πηγή πληροφοριών, εντούτοις, η πραγματική εικόνα μιας επιχείρησης αποτυπώνεται σε συνδυασμό με άλλες συμπληρωματικές πληροφορίες που περιλαμβάνονται στους ετήσιους απολογισμούς που καταρτίζουν οι επιχειρήσεις καθώς και σε εξωλογιστικά δεδομένα.

Οι βασικές χρηματοοικονομικές καταστάσεις των επιχειρήσεων είναι ο Ισολογισμός, η Κατάσταση Αποτελεσμάτων Χρήσης (ΚΑΧ), η Κατάσταση Ταμειακών Ροών και το Προσάρτημα ή Σημειώσεις. Σε αυτές παρουσιάζονται τα περιουσιακά στοιχεία, οι πηγές προέλευσης των κεφαλαίων τους, και τα οικονομικά αποτελέσματα από τις δραστηριότητές της καθώς και άλλες πληροφορίες που αφορούν τις χρησιμοποιούμενες λογιστικές μεθόδους κλπ. Όμως, οι συγκεκριμένες λογιστικές καταστάσεις παρουσιάζουν σημαντικά μειονεκτήματα, τα οποία είναι [1, 2]:

1. Τα στοιχεία που περιλαμβάνονται είναι πολύ συνοπτικά και ανομοιόμορφα καταταγμένα, καθώς δεν υπάρχει ενιαίος τύπος κατάρτισής τους για όλες τις επιχειρήσεις.
2. Το χρονικό διάστημα που μεσολαβεί μεταξύ του τέλους της χρήσεως, στην οποία αναφέρονται οι καταστάσεις, και της δημοσίευσής στο ευρύ κοινό είναι μεγάλος.

Για την επίλυση των προαναφερθέντων μειονεκτημάτων εφαρμόζεται η χρηματοοικονομική ανάλυση για τον υπολογισμό των κατάλληλων μεγεθών και σχέσεων που είναι σημαντικές και χρήσιμες για τη λήψη αποφάσεων οικονομικής φύσης. Συνεπώς, με άλλα λόγια η ανάλυση των λογιστικών καταστάσεων επιτυγχάνει τη μετατροπή πολλών και διαφορετικών στοιχείων, από απλούς αριθμούς σε χρήσιμες πληροφορίες, κάτι που δεν προσφέρεται αυτούσια.

Το είδος της χρηματοοικονομικής ανάλυσης των λογιστικών καταστάσεων των επιχειρήσεων εξαρτάται από το ιδιαίτερο ενδιαφέρον και τις επιδιώξεις αυτών που πραγματοποιούν την ανάλυση, οι οποίοι, συνήθως, είναι οι βραχυχρόνιοι πιστωτές, οι μακροχρόνιοι δανειστές, οι επενδυτές σε μετοχές της επιχείρησης, η διοίκηση και οι εργαζόμενοι

της επιχείρησης. Η ανάλυση των στοιχείων των λογιστικών καταστάσεων εξασφαλίζει ότι οι αποφάσεις των ενδιαφερομένων δεν στηρίζονται πλέον σε υποθέσεις, αλλά σε συγκεκριμένα αποτελέσματα, μειώνοντας την αβεβαιότητα στη λήψη αποφάσεων. Στη συνέχεια του κεφαλαίου αυτού θα παρουσιαστούν οι μέθοδοι ανάλυσης των λογιστικών καταστάσεων, ενώ θα δοθεί έμφαση στην ανάλυση με αριθμοδείκτες (ratio analysis).

2.2 Μέθοδοι Ανάλυσης Λογιστικών Καταστάσεων

Η χρηματοοικονομική ανάλυση των λογιστικών καταστάσεων στοχεύει στη μελέτη των σχέσεων των οικονομικών στοιχείων που αναφέρονται σε αυτές σε δεδομένη χρονική στιγμή, καθώς και των τάσεων αυτών, διαχρονικά. Η επίτευξη των αναλύσεων επιτυγχάνεται μέσω μεθόδων, οι οποίες μπορούν να ομαδοποιηθούν στις ακόλουθες κατηγορίες [3-5]:

1. Διαστρωματική ή κάθετη μέθοδος ανάλυσης (vertical analysis)
2. Συγκριτική ή διαχρονική μέθοδος ανάλυσης (comparative analysis)
3. Εξειδικευμένες μέθοδοι, που περιλαμβάνουν την ανάλυση των μεταβολών της οικονομικής θέσης μιας επιχείρησης, την ανάλυση νεκρού σημείου, καθώς και άλλες μέθοδοι.

2.2.1 Διαστρωματική ή Κάθετη Ανάλυση

Η διαστρωματική ή κάθετη ανάλυση περιλαμβάνει την κατάρτιση των χρηματοοικονομικών καταστάσεων του “κοινού μεγέθους” και τον υπολογισμό των κατάλληλων αριθμοδεικτών. Για τη διευκόλυνση των συγκρίσεων μεταξύ των στοιχείων μιας επιχείρησης για μια σειρά ετών καθώς και μεταξύ επιχειρήσεων, διαφορετικού μεγέθους, συντάσσονται οι ισολογισμοί και οι καταστάσεις αποτελεσμάτων χρήσης σε ποσοστά (%) του συνόλου των μεγεθών τους.

Στην ανάλυση “κοινού μεγέθους” κάθε στοιχείο του ισολογισμού διαιρείται με το σύνολο του ενεργητικού ή του παθητικού, ενώ τα στοιχεία της κατάστασης αποτελεσμάτων χρήσης διαιρούνται με το σύνολο των καθαρών πωλήσεων. Συνεπώς, η κατάσταση “κοινού μεγέθους” του ισολογισμού παρουσιάζει τα στοιχεία του σαν ποσοστό (%) του ενεργητικού ή παθητικού και η κατάσταση “κοινού μεγέθους” των αποτελεσμάτων χρήσης αποτυπώνουν τα στοιχεία αυτά σαν ποσοστό (%) των καθαρών πωλήσεων της χρήσης.

Πολλές φορές, η εύρεση του ποσοστού συμμετοχής κάθε στοιχείου στο σύνολο κάποιας κατηγορίας στοιχείων αποτελεί σημαντικό εργαλείο στη χρηματοοικονομική ανάλυση. Συγκεκριμένα, η έκφραση των απόλυτων μεγεθών σε ποσοστά είναι χρήσιμη στην ανάλυση της εσωτερικής διάρθρωσης των λογιστικών καταστάσεων και δείχνει τη σχετική σπουδαιότητα κάθε στοιχείου σε σχέση με το σύνολο των στοιχείων αυτών.

Η κάθετη ανάλυση του ισολογισμού εστιάζει σε δυο ουσιαστικά σημεία [5]:

1. Στον προσδιορισμό των πηγών προέλευσης των κεφαλαίων μιας επιχείρησης ή με άλλα λόγια στην κατανομή των κεφαλαίων της μεταξύ των ίδιων κεφαλαίων καθώς και των βραχυχρόνιων και μακροχρόνιων υποχρεώσεών της.
2. Στην κατανομή των κεφαλαίων της επιχείρησης μεταξύ των επιμέρους περιουσιακών στοιχείων της (κυκλοφορούν, πάγιο διαθέσιμα).

2.2.2 Συγκριτική ή Διαχρονική Ανάλυση

Η συγκριτική ή διαχρονική χρηματοοικονομική ανάλυση των λογιστικών καταστάσεων πραγματοποιείται με την κατάταξη των στοιχείων ανά κατηγορίες για μια σειρά ετών.

Τα λογιστικά στοιχεία αποκτούν ιδιαίτερη σημασία όταν αντιπαραβάλλονται με άλλα ίδια ή παρόμοια προηγούμενων ετών ή χρήσεων. Οι καταστάσεις με λογιστικά στοιχεία δύο ή περισσότερων χρήσεων καλούνται **συγκριτικές χρηματοοικονομικές καταστάσεις (comparative statements)**. Συνεπώς, ετήσια, εξαμηνιαία, τριμηνιαία ή μηνιαία στοιχεία μπορούν εύκολα να συγκριθούν με τα αντίστοιχα στοιχεία των προηγούμενων χρήσεων και να προκύψουν συμπεράσματα σχετικά με την τάση και το ρυθμό τους διαχρονικά.

Γίνεται σαφές ότι οι συγκριτικές χρηματοοικονομικές καταστάσεις ενισχύουν τη χρησιμότητά τους και εμφανίζουν με μεγαλύτερη σαφήνεια τη φύση και τις τάσεις των τρεχουσών μεταβολών που επηρεάζουν τη χρηματοοικονομική θέση των επιχειρήσεων. Μπορεί εύκολα να γίνει αντιληπτό, ότι όσο περισσότερες οι χρήσεις στις οποίες αναφέρονται οι συγκριτικές καταστάσεις, τόσο πληρέστερα και πιο αξιόπιστα είναι τα συμπεράσματα που μπορούν να εξαχθούν. Η εξασφάλιση της σύγκρισης μεταξύ των λογιστικών στοιχείων διαχρονικά, πρέπει οι συγκριτικές καταστάσεις να πληρούν τις ακόλουθες προϋποθέσεις [4, 6]:

1. Η ταξινόμηση των στοιχείων να είναι ομοιόμορφη σε όλες τις περιόδους που θα συμπεριληφθούν.
2. Στις ίδιες επιμέρους ομάδες στοιχείων να είναι ενταγμένα τα αυτά στοιχεία.
3. Να τηρούνται οι ίδιες λογιστικές αρχές σε όλη την περίοδο που αξιολογείται. Στην περίπτωση ύπαρξης αλλαγών θα πρέπει αυτές να σημειώνονται, ώστε να λαμβάνεται υπόψη από τον αναλυτή.
4. Να σημειώνονται οι ενδεχόμενες μεταβολές στις συνθήκες ή στη φύση των στοιχείων.

Στην περίπτωση που οι προαναφερθείσες προϋποθέσεις δεν τηρούνται, οι συγκρίσεις μπορούν να δώσουν παραπλανητική εικόνα να εξαχθούν εσφαλμένα συμπεράσματα. Οι συγκριτικές καταστάσεις είναι περισσότερο χρήσιμες αν παράλληλα με τα απόλυτα μεγέθη περιέχουν και ποσοστά [7].

Η συγκριτική ανάλυση περιλαμβάνει τις εξής μεθόδους:

1. **Συγκριτικές καταστάσεις συνήθους μορφής, διαχρονικά.** Οι καταστάσεις αυτές επιτρέπουν την αξιολόγηση των μεταβολών των διαφόρων ομάδων στοιχείων των λογιστικών καταστάσεων και των μεταβολών της χρηματοοικονομικής θέσης μιας επιχείρησης. Η απεικόνιση των μεταβολών αυτών γίνεται διαχρονικά τόσο σε απόλυτα μεγέθη όσο και σε ποσοστά. Η σύγκριση των στοιχείων των λογιστικών καταστάσεων για 2-3 χρόνια μπορεί να γίνει εύκολα με την παρακολούθηση των μεταβολών τόσο στα απόλυτα μεγέθη όσο και στα ποσοστά τους.

2. **Υπολογισμός αριθμοδεικτών τάσεως.** Για τη μελέτη στοιχείων μακράς χρονικής περιόδου (5-10 έτη), συνίσταται η σύγκριση των μεταβολών των διαφόρων μεγεθών με τον υπολογισμό των αριθμοδεικτών τάσεως. Ο υπολογισμός των συγκρινόμενων αριθμοδεικτών προϋποθέτει την επιλογή ενός έτους ή μιας χρήσης που θα αποτελέσει τη βάση αναφοράς. Οι αριθμοδείκτες τάσεως δίνουν πιο ανάγλυφα τις μεταβολές των οικονομικών στοιχείων διαχρονικά και καθιστούν δυνατή την οριζόντια ανάλυση και μελέτη αυτών.

3. **Διαχρονική ανάλυση των “κοινών μεγεθών” των λογιστικών καταστάσεων.** Η μέθοδος αυτή χρησιμοποιείται στην ανάλυση των χρονολογικών σειρών. Σύμφωνα με τη συγκεκριμένη τεχνική, γίνεται σύγκριση μεταξύ των διαφόρων κοινών μεγεθών μιας επιχείρησης διαχρονικά. Οι συγκρίσεις αυτές είναι πολύτιμες διότι δείχνουν τις μεταβολές των ποσοστών συμμετοχής των επί μέρους στοιχείων στους τομείς περιουσίας, υποχρεώσεων, κόστους και των άλλων κατηγοριών οικονομικών στοιχείων. Απαιτείται ιδιαίτερη προσοχή στην ερμηνεία των μεταβολών και των τάσεών τους διαχρονικά, διότι ένα ποσοστό μπορεί να μεταβληθεί είτε από μεταβολή του απόλυτου μεγέθους του αντίστοιχου στοιχείου είτε από τη μεταβολή του συνόλου του τομέα στον οποίο ανήκει. Για την ορθή ερμηνεία των καταστάσεων κοινών μεγεθών διαχρονικά, είναι απαραίτητο να εξεταστούν τόσο τα απόλυτα μεγέθη που αντιστοιχούν σε κάθε ποσοστό, όσο και η βάση πάνω στην οποία υπολογίστηκαν.

4. **Διαχρονική ανάλυση αριθμοδεικτών.** Η συνήθης χρήση των αριθμοδεικτών έγκειται στη σύγκρισή τους με τις τιμές των αντίστοιχων αριθμοδεικτών ομοειδών επιχειρήσεων σε ορισμένη χρονική στιγμή. Η ανάλυση αυτή ονομάζεται διαστρωματική και είναι στατικής μορφής, καθώς οι αριθμοδείκτες δεν είναι τίποτε άλλο παρά μια στιγμιαία απεικόνιση της χρηματοοικονομικής θέσης μιας επιχείρησης. Κύριος αντικειμενικός σκοπός της προαναφερθείσας ανάλυσης είναι η δημιουργία της απαραίτητης πληροφοριακής βάσης για τη λήψη ορθών αποφάσεων με τη σύγκριση ενός αριθμοδείκτη με τον αντίστοιχο αριθμοδείκτη άλλης ή άλλων επιχειρήσεων. Από την άλλη πλευρά, η εξέταση των αριθμοδεικτών μιας επιχείρησης διαχρονικά αναφέρεται στη συμπεριφορά αυτών για μια σειρά ετών και βοηθά στην πρόβλεψη της μελλοντικής χρηματοοικονομικής της κατάστασης. Επιπροσθέτως, μπορεί να διαπιστωθεί αν μια ευνοϊκή ή μια δυσμενής οικονομική κατάσταση μετατρέπεται σε μη ευνοϊκή ή βελτιώνεται με την πάροδο του χρόνου, αντίστοιχα. Χρειάζεται να επισημανθεί ότι για τη συγκεκριμένη ανάλυση απαιτείται η εφαρμογή στατιστικών μεθόδων για τη μελέτη μιας χρονικής σειράς στο παρελθόν όπως και στο μέλλον.

2.2.3 Εξειδικευμένες Μέθοδοι – Ανάλυση Νεκρού Σημείου

Υπάρχουν ειδικές τεχνικές που μπορούν να εφαρμοστούν για την ανάλυση των λογιστικών καταστάσεων των επιχειρήσεων. Στην παράγραφο αυτή θα παρουσιαστεί η ανάλυση νεκρού σημείου, σαν μια εξειδικευμένη μέθοδο αξιολόγησης των χρηματοοικονομικών καταστάσεων επιχειρήσεων.

Το **νεκρό σημείο** είναι το ποσό των πωλήσεων με το οποίο μια επιχείρηση καλύπτει ακριβώς τα σταθερά όσο και τα μεταβλητά της έξοδα, χωρίς να πραγματοποιεί ούτε κέρδος ούτε ζημιά. Το “νεκρό σημείο” αποτελεί σημαντικό αντικείμενο μελέτης και ανάγεται στη σφαίρα της ανάλυσης των πωλήσεων μιας επιχείρησης.

Η βασική αρχή στην οποία στηρίζεται η ανάλυση του “νεκρού σημείου” είναι η συμπεριφορά του κόστους. Αυτό οφείλεται στο γεγονός ότι ένα μέρος του κόστους είναι μεταβλητό και ανάλογο των πωλήσεων, ενώ ένα άλλο τμήμα του είναι σταθερό, τουλάχιστον για ένα μεγάλο εύρος πωλήσεων. Το σταθερό μέρος του κόστους αποτελείται από τις δαπάνες που παραμένουν αμετάβλητες και ανεξάρτητες από το ύψος και από το γεγονός της πραγματοποίησης ή όχι πωλήσεων. Δαπάνες που εντάσσονται σε αυτή την κατηγορία είναι τα έξοδα διοικήσεως, οι αποσβέσεις, τα ενοίκια γραφείων, μηχανών, τα χρηματοοικονομικά έξοδα και άλλα. Όμως, οι σταθερές δαπάνες μπορεί να μεταβάλλονται, αλλά η μεταβολή τους να είναι αποτέλεσμα άλλων αιτιών ανεξάρτητων από το μέγεθος της δραστηριότητας της επιχείρησης. Εξάλλου, είναι δυνατόν ορισμένες δαπάνες να παραμένουν αμετάβλητες μέχρι ενός ορισμένου ύψους πωλήσεων, πέραν του οποίου απαιτούνται πρόσθετες δαπάνες.

Από την άλλη πλευρά, οι μεταβλητές δαπάνες είναι ανάλογες προς το ύψος των πωλήσεων ή τον κύκλο εργασιών της επιχείρησης. Εκτός από τις παραπάνω κατηγορίες δαπανών, υπάρχουν και οι ημιμεταβλητές δαπάνες που διαχωρίζονται στο σταθερό και το μεταβλητό τους τμήμα, όπως υπάρχουν και άλλες ημιμεταβλητές δαπάνες που δεν μεταβάλλονται ανάλογα με το ύψος των πωλήσεων. Συνεπώς, ουσιαστικά δίνεται έμφαση μόνο σε δυο είδη δαπανών, τις σταθερές και τις μεταβλητές.

Η ανάλυση του “νεκρού σημείου” αποτελεί τη βασική μέθοδο μελέτης και προγραμματισμού των πωλήσεων με βάση την υπάρχουσα σχέση μεταξύ των σταθερών και των μεταβλητών δαπανών της επιχείρησης. Η συγκεκριμένη ανάλυση θα πρέπει να θεωρηθεί σαν ο οδηγός για τη λήψη ορθολογικών επιχειρηματικών αποφάσεων. Ο υπολογισμός του “νεκρού σημείου” δείχνει το ελάχιστο ύψος πωλήσεων, που πρέπει να πραγματοποιεί μια επιχείρηση για να καλύπτονται τόσο οι σταθερές όσο και οι μεταβλητές δαπάνες της. Με άλλα λόγια, προσδιορισμός του “νεκρού σημείου” καθορίζει το βαθμό περιορισμού των πωλήσεων της επιχείρησης, χωρίς αυτή να παρουσιάζει κέρδη ή ζημίες.

Είναι προφανές, ότι κάθε επιχείρηση έχει το δικό της “νεκρό σημείο” και είναι εκείνο στο οποίο οι πωλήσεις της ισούνται με το συνολικό κόστος παραγωγής των προϊόντων της, οπότε το

οικονομικό της αποτέλεσμα είναι μηδέν. Αν οι πωλήσεις της επιχείρησης είναι μεγαλύτερες από αυτές του “νεκρού σημείου”, τότε η επιχείρηση πραγματοποιεί κέρδος, ενώ στην περίπτωση που είναι χαμηλότερες, τότε η επιχείρηση έχει ζημιά.

Ο προσδιορισμός και η ανάλυση του “νεκρού σημείου” είναι πολύ χρήσιμα μέσα στις εξής περιπτώσεις [2, 5, 8]:

1. Στην πρόβλεψη, του πιθανού κατά μονάδα κόστους προϊόντων, στα διάφορα επίπεδα παραγωγής του και του αναγκαίου ύψους πωλήσεων, οι οποίες θα δικαιολογούν μια προγραμματισμένη επένδυση, που θα οδηγεί στην αύξηση της παραγωγικής δυναμικότητας της επιχείρησης και της πιθανής επίδρασης των κερδών από μια μεταβολή στο κόστος των πρώτων υλών ή των αμοιβών εργασίας ή μια μεταβολή στον όγκο ή τις τιμές πωλήσεων των προϊόντων της επιχείρησης. Σχετικά με τις μεταβολές των τιμών πωλήσεων, η επίδρασή τους στον όγκο πωλήσεων θα πρέπει να εξετάζεται σε συνδυασμό με το βαθμό ελαστικότητας της ζήτησης των προϊόντων αυτής.

2. Στην αξιολόγηση της ικανότητας της διοικήσεως της επιχείρησης. Αυτό επιτυγχάνεται με τη σύγκριση και αξιολόγηση του προγραμματικά επιτευχθέντος “νεκρού σημείου” με τα προκαθορισμένα επίπεδα αυτού.

3. Για τη λήψη επιχειρηματικών αποφάσεων σχετικά με τις τρέχουσες δραστηριότητες της επιχείρησης, την πολιτική που ακολουθεί και άλλα ζητήματα.

Οι σημαντικότερες παραδοχές, που γίνονται συνήθως στην πράξη, προκειμένου να υπολογιστεί το “νεκρό σημείο” είναι οι ακόλουθες [5]:

1. Το συνολικό κόστος και τα έσοδα μπορούν να προσδιοριστούν με αξιοπιστία και η μεταξύ τους σχέση είναι γραμμική για ένα ορισμένο εύρος παραγωγής.

2. Το κόστος μπορεί χωρίς δυσκολία να διαχωριστεί στο σταθερό και το μεταβλητό του μέρος, χωρίς το σταθερό να μεταβάλλεται και το μεταβλητό να είναι ανάλογο με το ύψος των πωλήσεων.

3. Οι τιμές πώλησης των προϊόντων της επιχείρησης και οι τιμές των συντελεστών παραγωγής δεν μεταβάλλονται.

4. Δεν υπήρξε σημαντική μεταβολή του γενικού επιπέδου των τιμών, κατά την περίοδο που ζητείται ο προσδιορισμός του “νεκρού σημείου”.

5. Οι μέθοδοι παραγωγής και η παραγωγική δυναμικότητα της επιχείρησης δεν μεταβάλλονται.

6. Ο μόνος συντελεστής που επηρεάζει το κόστος είναι το ύψος των πωλήσεων.

2.3 Ανάλυση Λογιστικών Καταστάσεων με Χρήση Αριθμοδεικτών

Η χρήση *αριθμοδεικτών (ratio analysis)* αποτελεί μια από τις πιο διαδεδομένες και δυναμικές μεθόδους χρηματοοικονομικής ανάλυσης. Οι μέθοδοι που έχουν παρουσιαστεί συμπληρώνονται με τη χρησιμοποίηση των αριθμοδεικτών, που βοηθούν στην ερμηνεία των οικονομικών στοιχείων των επιχειρήσεων.

Ο *αριθμοδείκτης* είναι η απλή σχέση ενός στοιχείου του ισολογισμού ή της κατάστασης αποτελεσμάτων χρήσης προς ένα άλλο στοιχείο. Ο λόγος χρησιμοποίησης των αριθμοδεικτών έγκειται στην απαίτηση για εύρεση ενός εργαλείου ώστε να γίνεται αμέσως αντιληπτή η πραγματική αξία και η σπουδαιότητα των απόλυτων μεγεθών. Ο αριθμοδείκτης για να έχει αξία, θα πρέπει να εκφράζει σχέση που παρουσιάζει οικονομικό ενδιαφέρον και οδηγεί σε συγκεκριμένα συμπεράσματα. Κρίνεται σκόπιμο να επισημανθεί ότι ο υπολογισμός και η παρουσίαση των διαφόρων αριθμοδεικτών αποτελεί μια μέθοδο ανάλυσης που πολλές φορές παρέχει μόνο ενδείξεις. Συνεπώς, ένας μεμονωμένος αριθμοδείκτης δεν μπορεί να δώσει πλήρη εικόνα της οικονομικής θέσης μιας επιχείρησης αν δεν συσχετιστεί με τους αντίστοιχους αριθμοδείκτες μιας σειράς προηγούμενων χρήσεων.

Στη βιβλιογραφία υπάρχουν πολλοί διαφορετικοί αριθμοδείκτες [2-6, 8, 9]. Οι κυριότεροι και οι περισσότερο χρησιμοποιούμενοι αριθμοδείκτες στη χρηματοοικονομική ανάλυση των λογιστικών καταστάσεων μπορούν να ομαδοποιηθούν στις ακόλουθες κατηγορίες [2, 3, 6, 8, 10]:

1. Αριθμοδείκτες ρευστότητας (liquidity ratios).
2. Αριθμοδείκτες δραστηριότητας (activity ratios).
3. Αριθμοδείκτες αποδοτικότητας (profitability ratios).
4. Αριθμοδείκτες διάρθρωσης κεφαλαίων και βιωσιμότητας (financial structure and viability ratios).
5. Αριθμοδείκτες επενδύσεων ή επενδυτικοί αριθμοδείκτες (investment ratios).

2.3.1 Αριθμοδείκτες Ρευστότητας

Οι *αριθμοδείκτες ρευστότητας (liquidity ratios)* χρησιμοποιούνται για τον προσδιορισμό τόσο της βραχυχρόνιας οικονομικής θέσης μιας επιχείρησης όσο και της ικανότητά της να ανταποκριθεί στις βραχυχρόνιες υποχρεώσεις της [2, 4, 8].

Η ύπαρξη ρευστότητας σε μια επιχείρηση έχει επίδραση στα κέρδη της, διότι αν τα κυκλοφορούντα στοιχεία της δεν μετατρέπονται εύκολα σε μετρητά ή αν δεν υπάρχουν επαρκή μετρητά διαθέσιμα, τότε η επιχείρηση δεν είναι σε θέση να πληρώσει τις υποχρεώσεις της. Η τοποθέτηση κεφαλαίων σε κυκλοφορούντα στοιχεία δεν αποτελεί κατά κανόνα παραγωγική επένδυση και ως εκ τούτου η διατήρηση υπερβολικά υψηλών κυκλοφορούντων στοιχείων ενδέχεται να μειώσει την κερδοφόρα δυναμικότητα της επιχείρησης. Στο ίδιο αποτέλεσμα μπορεί να οδηγήσει και η ανεπάρκεια κυκλοφορούντων στοιχείων, διότι η αδυναμία κάλυψης των

υποχρεώσεων καταλήγει σε οικονομική αστάθεια με καταστρεπτικές συνέπειες για την επιχείρηση.

Οι αριθμοδείκτες ρευστότητας προορίζονται να δώσουν την εικόνα της τρέχουσας οικονομικής κατάστασης μιας επιχείρησης. Ο προσδιορισμός της τρέχουσας οικονομικής κατάστασης ενδιαφέρει τόσο τη διοίκησή της όσο και τους εκτός αυτής πιστωτές και μετόχους της. Ο υπολογισμός των αριθμοδεικτών ρευστότητας επιτρέπει τον έλεγχο του τρόπου απασχόλησης των κεφαλαίων κίνησης μέσα στην επιχείρηση. Επιπρόσθετα, ελέγχεται η επάρκεια των κεφαλαίων σε σύγκριση με τις εργασίες της ή αν υπολείπονται από τα κανονικά οπότε υπάρχει πρόβλημα ρευστών. Η διαχρονική παρακολούθηση των αριθμοδεικτών ρευστότητας παρέχει τη δυνατότητα διαπίστωσης βελτίωσης ή επιδείνωσης της τρέχουσας οικονομικής κατάστασης της επιχείρησης.

Σε γενικές γραμμές για να θεωρηθεί ότι μια επιχείρηση βρίσκεται σε καλή οικονομική κατάσταση από άποψης κεφαλαίων κίνησης θα πρέπει να πληρούνται οι παρακάτω προϋποθέσεις [2, 5]:

1. Να μπορεί να ανταποκρίνεται στις καθημερινές απαιτήσεις των βραχυχρόνιων δαπανών της και να είναι σε θέση να εξοφλεί τις ληξιπρόθεσμες υποχρεώσεις της.
2. Να μπορεί να εκτελεί απρόσκοπτα τις εργασίες της και να μπορεί να εκμεταλλεύεται ευκαιρίες που ενδεχομένως να παρουσιαστούν.
3. Να μπορεί να ανταποκρίνεται στην εξόφληση των δεδουλευμένων τόκων και μερισμάτων των μετόχων της.

2.3.2 Αριθμοδείκτες Δραστηριότητας

Οι *αριθμοδείκτες δραστηριότητας (activity ratios)* προσδιορίζουν το βαθμό μετατροπής ορισμένων περιουσιακών στοιχείων (αποθεμάτων, απαιτήσεων) σε μετρητά. Με άλλα λόγια, οι συγκεκριμένοι αριθμοδείκτες μετρούν την παραγωγικότητα και το βαθμό χρησιμοποίησης των περιουσιακών στοιχείων της επιχείρησης. Γενικά, όσο πιο εντατική είναι η χρήση των στοιχείων μιας επιχείρησης τόσο τούτο αποβαίνει προς όφελός της.

Η μείωση του αριθμοδείκτη ταχύτητας είσπραξης απαιτήσεων ενδέχεται να αντανakλά τα ακόλουθα [9, 14]:

1. Μείωση των πωλήσεων με παράλληλη αύξηση των απαιτήσεων.
2. Μείωση των απαιτήσεων και αναλογικά μεγαλύτερη μείωση των πωλήσεων.
3. Αύξηση των πωλήσεων και αναλογικά μεγαλύτερη αύξηση των απαιτήσεων.
4. Μείωση των πωλήσεων χωρίς ταυτόχρονη μείωση των απαιτήσεων.
5. Αύξηση των απαιτήσεων χωρίς ταυτόχρονη μεταβολή των πωλήσεων.

2.3.3 Αριθμοδείκτες Αποδοτικότητας

Οι *αριθμοδείκτες αποδοτικότητας (profitability ratios)* αναφέρονται στη σχέση μεταξύ των κερδών και των απασχολούμενων κεφαλαίων της επιχείρησης καθώς και στη σχέση μεταξύ κερδών και πωλήσεων. Η σχέση μεταξύ των κερδών και των απασχολούμενων κεφαλαίων είναι ένας από τους πιο συχνά χρησιμοποιούμενους δείκτες μέτρησης της απόδοσης μιας επιχείρησης, γιατί επιτρέπει την σύγκριση με την απόδοση άλλων εναλλακτικών μορφών επενδύσεων καθώς και με την απόδοση άλλων επιχειρήσεων.

2.3.4 Αριθμοδείκτες Δαπανών Λειτουργίας

Οι *αριθμοδείκτες δαπανών λειτουργίας (operating expense ratios)* παρέχουν ένδειξη της πολιτικής που ακολουθεί η διοίκηση μιας επιχείρησης έναντι των διαφόρων δαπανών λειτουργίας της και της αποτελεσματικότητας έναντι των επιμέρους δαπανών. Μια ανοδική πορεία λειτουργικών δαπανών πρέπει να διερευνηθεί μήπως δεν καταβάλλονται προσπάθειες συμπίεσης αυτών ή αυξάνονται δυσανάλογα των πωλήσεων ή των άλλων μεγεθών της επιχείρησης.

2.3.5 Αριθμοδείκτες Διάρθρωσης Κεφαλαίων και Βιωσιμότητας

Οι αριθμοδείκτες που αναπτύχθηκαν στα προηγούμενα εδάφια εξέταζαν τις βραχυπρόθεσμες δυνατότητες μιας επιχείρησης. Με άλλα λόγια μελετήθηκε η ικανότητα της επιχείρησης να ανταποκρίνεται στην εξόφληση των τρεχουσών υποχρεώσεων της όταν αυτές γίνουν απαιτητές. Από την άλλη πλευρά ιδιαίτερη βαρύτητα έχει η μελέτη της μακροχρόνιας οικονομικής κατάστασης, η οποία διαφέρει σημαντικά από την βραχυχρόνια εκτίμηση. Η διαφοροποίηση αυτή έγκειται στο γεγονός ότι τα χρονικά όρια είναι μικρά στην ανάλυση της βραχυχρόνιας κατάστασης, με αποτέλεσμα να είναι δυνατή η πρόβλεψη των στοιχείων εκείνων που την προσδιορίζουν. Από την άλλη πλευρά, δεν συμβαίνει το ίδιο για μια μακρά χρονική περίοδο.

Η διαδικασία μελέτης της οικονομικής κατάστασης μιας επιχείρησης σε μακροχρόνιο επίπεδο, περιλαμβάνει την ανάλυση της διάρθρωσης των κεφαλαίων της. Ο όρος διάρθρωση των κεφαλαίων μιας επιχείρησης αναφέρεται στα διάφορα είδη και τις μορφές των κεφαλαίων που χρησιμοποιεί για τη χρηματοδότησή της. Συγκεκριμένα, υπάρχουν τα ίδια κεφάλαια, καθώς και τις βραχυπρόθεσμες, μεσοπρόθεσμες και μακροπρόθεσμες υποχρεώσεις της. Οι διάφορες μορφές υποχρεώσεων μπορούν να εξασφαλιστούν είτε με την παροχή υποθήκης ή προσημείωσης ή ενεχύρου είτε χωρίς την παροχή εμπράγματης ασφάλειας στους πιστωτές της. Γίνεται, επομένως, σαφές ότι οι διάφορες μορφές δανειακών κεφαλαίων μιας επιχείρησης περιέχουν διάφορα ποσοστά κινδύνου για τους πιστωτές της.

Η σπουδαιότητα της διάρθρωσης των κεφαλαίων προέρχεται από την ουσιαστική διαφορά που υπάρχει μεταξύ των ίδιων και των δανειακών κεφαλαίων. Τα ίδια κεφάλαια είναι εκείνα που επωμίζονται τον επιχειρηματικό κίνδυνο που αναπόφευκτα υπάρχει σε κάθε επιχείρηση. Το χαρακτηριστικό των ίδιων κεφαλαίων είναι ότι δεν έχουν ορισμένο χρόνο επιστροφής ούτε εξασφαλισμένη απόδοση, δεδομένου ότι η διανομή μερίσματος στους μετόχους εξαρτάται από το ύψος των κερδών και την απόφαση της διοίκησης και της γενικής συνέλευσης των μετόχων. Τα ίδια κεφάλαια, επειδή θεωρούνται μόνιμα, επενδύονται κατά κανόνα σε μακροχρόνιες επενδύσεις και εκτίθεται σε περισσότερους κινδύνους από ότι τα δανειακά κεφάλαια.

Τα δανειακά κεφάλαια πρέπει να εξοφλούνται προσαυξημένα με τους τόκους τους σε τακτά χρονικά διαστήματα, ανεξάρτητα από την οικονομική θέση της επιχείρησης. Σε περίπτωση που μια επιχείρηση δεν είναι σε θέση να εξοφλήσει τις υποχρεώσεις, που απορρέουν από τη χρησιμοποίηση των ξένων κεφαλαίων (επιστροφή κεφαλαίων και πληρωμή τόκων), θα έχει σαν συνέπεια να υποστούν ζημιά τα ίδια κεφάλαιά της. Όσο μεγαλύτερη είναι η αναλογία των ξένων κεφαλαίων προς το σύνολο των κεφαλαίων της επιχείρησης, τόσο περισσότερο χρεωμένη εμφανίζεται αυτή και τόσο μεγαλύτερες είναι οι σταθερές επιβαρύνσεις και υποχρεώσεις της για την εξόφλησή τους. Με άλλα λόγια, υπάρχει μεγάλη πίεση στην επιχείρηση για την πληρωμή τόκων και την επιστροφή των δανειακών κεφαλαίων, όταν αυτά καταστούν ληξιπρόθεσμα. Η πίεση αυτή είναι περισσότερο αισθητή σε περίπτωση που σημειωθεί κάμψη των πωλήσεων και των κερδών της επιχείρησης.

Όταν τα κέρδη μιας επιχείρησης παρουσιάζουν διακυμάνσεις, τότε η ύπαρξη μεγάλου ποσοστού δανειακών κεφαλαίων τείνει να αυξήσει το δανεισμό, που ενδεχομένως να επιδράσει στα κέρδη και να τα καταστήσει περισσότερο ασταθή. Εξάλλου, μεγάλο ποσοστό υποχρεώσεων στη διάρθρωση των κεφαλαίων μιας επιχείρησης αυξάνει τον κίνδυνο τόσο για τους επενδυτές μετόχους όσο και για τους πιστωτές της. Αυτό οφείλεται στο γεγονός ότι αν οι υποχρεώσεις προς τους πιστωτές δεν ικανοποιηθούν, τότε θα επιδιώξουν την εξόφλησή τους με οποιοδήποτε τρόπο, οδηγώντας την επιχείρηση ακόμη και σε πτώχευση.

Η χρησιμοποίηση ξένων κεφαλαίων για χρηματοδότηση προγραμμάτων νέων επενδύσεων παρουσιάζει πολλά πλεονεκτήματα σε σχέση με την έκδοση νέων τίτλων από αύξηση κεφαλαίου. Στο παρελθόν παρατηρούνταν μια τάση των επιχειρήσεων προς το δανεισμό για τη χρηματοδότησή τους, λόγω των υψηλών ρυθμών πληθωρισμού που υπήρχαν. Η τάση αυτή είναι κατά κανόνα περισσότερο έντονη σε περιόδους πληθωρισμού και μείωσης της αγοραστικής αξίας του χρήματος, διότι τότε τα ονομαστικά κέρδη των επιχειρήσεων εμφανίζονται αυξημένα ενώ οι υποχρεώσεις τους για την εξυπηρέτηση των δανειακών κεφαλαίων παραμένουν σταθερές. Επιπλέον, οι τόκοι των ξένων κεφαλαίων αφαιρούνται από τα κέρδη και επομένως, καταβάλλονται λιγότεροι φόροι.

Σχετικά με την επίδραση που έχει η διάρθρωση των κεφαλαίων μιας επιχείρησης στην αξία των μετοχών της και στο ύψος των κερδών της υπάρχουν δυο διαφορετικές προσεγγίσεις. Η μια

προσέγγιση υποστηρίζει ότι τα δανειακά κεφάλαια, εκτός από το όφελος που αποφέρουν στην επιχείρηση λόγω των φορολογικών ελαφρύνσεων, ασκούν επίσης, επίδραση στα κέρδη και στην εκτίμηση της αξίας των μετοχών της. Η δεύτερη θεώρηση ισχυρίζεται ότι εκτός από το φορολογικό όφελος, τα δανειακά κεφάλαια δεν επηρεάζουν τα κέρδη και την εκτίμηση της αξίας των μετοχών της.

Η άριστη διάρθρωση κεφαλαίων για μια επιχείρηση είναι εκείνη που κατορθώνει να επιφέρει ισορροπία μεταξύ του αναλαμβανόμενου κινδύνου και της απόδοσης της επιχείρησης και η οποία μεγιστοποιεί την τιμή των μετοχών της στην Αγορά, ενώ ταυτόχρονα ελαχιστοποιεί το κόστος των κεφαλαίων της. Εφόσον η μεγιστοποίηση της τιμής των μετοχών ταυτίζονται χρονικά, το πρόβλημα εύρεσης της άριστης διάρθρωσης κεφαλαίων είναι πρόβλημα μεγιστοποίησης της αξίας της ή ελαχιστοποίησης του κόστους των κεφαλαίων της.

Γενικότερα, οι παράγοντες που επηρεάζουν τις αποφάσεις μιας επιχείρησης για τη διατήρηση δεδομένης διάρθρωσης κεφαλαίων είναι οι ακόλουθοι [3, 17]:

1. Ο **επιχειρηματικός κίνδυνος** που υπάρχει σε όλες τις οικονομικές δραστηριότητες και συνδέεται άμεσα με το είδος της κάθε επιχείρησης. Όσο μεγαλύτερος είναι ο επιχειρηματικός κίνδυνος, τόσο χαμηλότερος πρέπει να είναι ο αριθμοδείκτης δανειακής επιβάρυνσης.

2. Η **θέση της επιχείρησης από άποψη φορολογίας**. Είναι γνωστό ότι οι τόκοι των ξένων κεφαλαίων αφαιρούνται από τα κέρδη των επιχειρήσεων, με αποτέλεσμα μείωση του φορολογητέου ποσού και του κόστους δανεισμού ανάλογα με το συντελεστή φορολόγησης κάθε επιχείρησης. Το πλεονέκτημα αυτό είναι άνευ σημασίας, αν το μεγαλύτερο μέρος των κερδών της επιχείρησης έχεις λόγους να υπόκειται σε φορολογικές απαλλαγές. Αυτό παρατηρείται όταν η επιχείρηση έχει τη δυνατότητα να πραγματοποιεί αυξημένες αποσβέσεις ή να έχει μειωμένη φορολογική επιβάρυνση λόγω πραγματοποίησης νέων επενδύσεων ή τέλος να απαλλάσσεται εντελώς από φόρους, λόγω μεταφοράς ζημιών προηγούμενων χρήσεων. Στις προαναφερθείσες περιπτώσεις ο συντελεστής φορολογίας των κερδών της επιχείρησης είναι πολύ χαμηλός ή και μηδενικός, οπότε το κίνητρο των φορολογικών ελαφρύνσεων από τη χρησιμοποίηση κεφαλαίων ατονεί.

3. Η **ικανότητα της επιχείρησης να αντλεί κεφάλαια με επωφελείς για αυτήν όρους, ακόμη και κάτω από δύσκολες συνθήκες**. Επιχειρήσεις με αμετάβλητο κύκλο εργασιών έχουν ανάγκη συνεχούς ροής κεφαλαίων για την επίτευξη μακροχρόνιας ανόδου. Κάτω από συνθήκες μειωμένης ρευστότητας αυξάνονται οι δυσκολίες δανεισμού από τις τράπεζες, διότι οι τελευταίες επιλέγουν με αυστηρά κριτήρια τις επιχειρήσεις που δανείζουν και προτιμούν αυτές που παρουσιάζουν υγιείς και δυναμικούς ισολογισμούς. Συνεπώς, οι ανάγκες μιας επιχείρησης για κεφάλαια σε μελλοντικό χρόνο και οι συνέπειες από τυχόν έλλειψή τους έχουν σημαντική επίδραση στη διάρθρωση των κεφαλαίων της. Η ανάλυση μιας επιχείρησης σε οικονομικό επίπεδο, απαιτεί τον προσδιορισμό της θέσης της σε βραχυχρόνια και μακροχρόνια σκοπιά. Συγκεκριμένα, μια ευνοϊκή τρέχουσα οικονομική κατάσταση μπορεί να εξουδετερωθεί από μια

αντίθετη κατάσταση στην περιουσιακή και κεφαλαιακή της διάρθρωση. Από την άλλη πλευρά, μια ευνοϊκή οικονομική κατάσταση από μακροχρόνια σκοπιά, μπορεί να εξουδετερωθεί από μια δυσμενή τρέχουσα κατάσταση στα κεφάλαια κίνησης της επιχείρησης.

2.3.6 Επενδυτικοί Αριθμοδείκτες

Οι *επενδυτικοί αριθμοδείκτες (investment ratios)* χρησιμοποιούνται από τους επενδυτές για τη λήψη αποφάσεων αναφορικά με την αγορά, πώληση ή διατήρηση μιας επένδυσης σε μετοχικούς τίτλους μιας επιχείρησης.

Οι περισσότεροι από τους ευρύτατα χρησιμοποιούμενους επενδυτικούς αριθμοδείκτες συσχετίζουν την τρέχουσα τιμή των μετοχών μιας επιχείρησης με ανά μετοχή μεγέθη του ισολογισμού και της κατάστασης αποτελεσμάτων χρήσης. Η αναφορά σε ανά μετοχή σχέσεις είναι πολύ χρήσιμη από πρακτικής πλευράς, διότι οι τιμές που αναφέρονται στη χρηματιστηριακή αγορά αφορούν μια μετοχή και δεν παρουσιάζουν συνολικά μεγέθη της επιχείρησης.

2.4 Περιορισμοί & Αδυναμίες της Ανάλυσης με Χρηματοοικονομικούς Δείκτες

Πριν την παρουσίαση των Τεχνητών Νευρωνικών Δικτύων, στο επόμενο κεφάλαιο, είναι σκόπιμο να επισημανθούν οι αδυναμίες που παρουσιάζουν οι δημοσιευμένες λογιστικές καταστάσεις και η ανάλυση με χρηματοοικονομικούς δείκτες. Παρόλο που η ανάλυση με χρηματοοικονομικούς δείκτες αποτελεί ένα ισχυρό εργαλείο, εντούτοις έχει ορισμένες σημαντικές αδυναμίες [17]. Οι αναλυτές οφείλουν να είναι προσεκτικοί και να μην εφαρμόζουν τις τεχνικές χρηματοοικονομικής ανάλυσης σε οποιοδήποτε σύνολο ισολογισμών και καταστάσεων αποτελεσμάτων χρήσης. Στη συνέχεια του εδαφίου παρουσιάζονται ορισμένα από τα πιο σημαντικά προβλήματα [2, 5].

Σε εταιρείες και μεγάλες επιχειρήσεις πολλαπλών δραστηριοτήτων είναι πιθανό να παρουσιαστεί εντονότερα το πρόβλημα κατά την χρηματοοικονομική ανάλυση. Οι εταιρείες αυτές συχνά έχουν τμήματα που λειτουργούν σε σημαντικά διαφορετικές βιομηχανίες. Οι οικονομικές πληροφορίες, που δημοσιεύουν οι επιχειρήσεις, περιέχουν ενοποιημένα τα αποτελέσματα των διαφόρων αυτών εργασιών σε ένα συνολικό ισολογισμό. Η ερμηνεία των χρηματοοικονομικών δεικτών εξαρτάται σε μεγάλο βαθμό από τα πρότυπα της επιχείρησης, με αποτέλεσμα η ανάμιξη των αποτελεσμάτων από διαφορετικές επιχειρήσεις να μπορούν να μειώσουν σημαντικά την ενημερωτική αξία της χρηματοοικονομικής ανάλυσης.

Οι αναφορές χρηματοοικονομικής ανάλυσης που καθορίζονται από τους ειδικούς οικονομολόγους για επιχειρήσεις που δραστηριοποιούνται σε πολλούς τομείς απαιτούν τη

δημοσιοποίηση επιπρόσθετων πληροφοριών ανά τομέα δραστηριοποίησης, αλλά γενικά αποτελεί περιορισμένο πεδίο εφαρμογής και χρησιμοποίησης [17].

Οι χρηματοοικονομικές αρχές παρέχουν μεγάλη διακριτική ευχέρεια για την υποβολή εκθέσεων. Αυτό σημαίνει ότι παρόμοιες εταιρείες μπορούν να παρουσιάσουν τα ίδια στοιχεία με διαφορετικό τρόπο, με συνέπεια τα οικονομικά τους αποτελέσματα να είναι τεχνητά ανόμοια. Η απόσβεση αποτελεί ένα τέτοιο στοιχείο. Η επιλογή μεταξύ της ταχείας και της σταθερής απόσβεσης επαφίεται στην επιχείρηση, αλλά η διαφορά μπορεί να διπλασιάσει την αναφερθείσα απόσβεση σε μια δεδομένη χρονική περίοδο. Αυτό με τη σειρά του μπορεί να κάνει μια μεγάλη ποσοστιαία διαφορά του καθαρού εισοδήματος μεταξύ των δύο κατ' ουσία όμοιων επιχειρήσεων.

Ο ισολογισμός παρουσιάζει συνήθως τα διάφορα μεγέθη της επιχείρησης σε παρελθούσες και όχι τρέχουσες τιμές. Επομένως, όλα τα περιουσιακά στοιχεία και υποχρεώσεις της εμφανίζονται συνήθως σε τιμές ιστορικού κόστους βασισμένες σε παρελθοντικά γεγονότα. Αυτό συνεπάγεται την απεικόνιση των μεγεθών του ισολογισμού σε τιμές που αντιπροσωπεύουν τις αξίες κατά το χρόνο που συνέβησαν τα αντίστοιχα λογιστικά γεγονότα. Ο πληθωρισμός στρεβλώνει συχνά τις οικονομικές καταστάσεις καθώς οι τιμές μεταβάλλονται σημαντικά από έτος σε έτος [5, 17]. Αυτό έχει σαν συνέπεια να απαιτείται η αναπροσαρμογή των ιστορικών τιμών των μεγεθών του ισολογισμού, για να αντισταθμίζεται η απώλεια της αγοραστικής αξίας του νομίσματος στο οποίο αναγράφονται τα ποσά. Όμως, αυτή η διαδικασία δεν γίνεται στην πράξη, με αποτέλεσμα ο ισολογισμός να απεικονίζει στοιχεία με αξίες του ίδιου νομίσματος, αλλά με διαφορετική αγοραστική αξία μεταξύ των ετών. Ειδικότερα, ακίνητα που αγοράστηκαν πριν από χρόνια, θα καταχωρηθούν στον ισολογισμό στο αρχικό κόστος τους. Ωστόσο, τα ακίνητα μπορεί να αξίζουν σήμερα πολύ περισσότερο από την τιμή αγοράς τους. Κατά τη διάρκεια περιόδων μεγάλου πληθωρισμού, η απογραφή, το κόστος των πωληθέντων αγαθών και οι αποσβέσεις μπορεί να στρεβλώσουν τα πραγματικά αποτελέσματα. Η ερμηνεία των χρηματοοικονομικών δεικτών δεν είναι πάντοτε σαφής.

Υπάρχει ακόμα μια άλλη οπτική γωνία που μπορεί να θεωρηθεί ότι ο ισολογισμός παρουσιάζει αδυναμίες και ανεπάρκειες. Υπάρχουν στοιχεία που είναι αδύνατο να μετρηθούν και να απεικονιστούν, όπως η φήμη και η πελατεία της επιχείρησης που έχουν αποκτηθεί από τη μακροχρόνια λειτουργία της, καθώς και άλλα άυλα στοιχεία. Η μη απεικόνιση των στοιχείων αυτών είναι σύμφωνη με την αρχή της αντικειμενικότητας, από την άλλη όμως, αποτελεί ελλιπή πληροφόρηση που ενδέχεται να επηρεάζει τις αποφάσεις των ενδιαφερομένων.

Ένα επιπρόσθετο στοιχείο αδυναμίας των λογιστικών καταστάσεων εστιάζει στα έξοδα που πραγματοποιήθηκαν σε μια χρήση με σκοπό να υπάρξει μελλοντική απόδοση. Στην κατηγορία αυτή εντάσσονται τα έξοδα οργάνωσης και πρώτης εγκατάστασης καθώς και τα έξοδα εκπαίδευσης του προσωπικού.

Δεν μπορούν να παραλειφθούν ένα πλήθος παραγόντων που ενώ δεν αποτυπώνονται στις λογιστικές καταστάσεις μιας επιχείρησης, εντούτοις, τις επηρεάζουν δυσμενώς. Συγκεκριμένα, η

αλλαγή επιθυμιών της αγοράς, οι μεταρρυθμίσεις στη φορολογική νομοθεσία και η αλλαγή της εμπορικής πολιτικής της χώρας που είναι εγκατεστημένη η επιχείρηση ή των χωρών με τις οποίες συναλλάσσεται είναι μερικές παράμετροι που επιδρούν στα στοιχεία των λογιστικών καταστάσεων.

Ωστόσο πέρα των προαναφερθέντων αδυναμιών, η ανάλυση με χρηματοοικονομικούς δείκτες είναι πολύ σημαντική τόσο για τους επενδυτές όσο και για τους πιστωτές της επιχείρησης. Το πιο σπουδαίο στοιχείο που χρειάζεται να επισημανθεί είναι ότι η ανάλυση που βασίζεται στους χρηματοοικονομικούς δείκτες δεν μπορεί να δώσει μόνη της απαντήσεις, αλλά πρέπει να δοθεί έμφαση στις πρόσθετες πληροφορίες που παρουσιάζονται με την μορφή υποσημειώσεων ή διευκρινήσεων, οι οποίες συνοδεύουν τους δημοσιευμένους ισολογισμούς [17].

ΒΙΒΛΙΟΓΡΑΦΙΑ ΚΕΦΑΛΑΙΟΥ

1. Berk J., DeMarzo P., and Harford J., “Fundamentals of Corporate Finance”, 2nd edition, *Prentice Hall*, 2012 (ISBN 978-0-13-214823-8).
2. Ross S., Westerfield R., and Jordan B., “Fundamentals of Corporate Finance”, 8th edition, *McGraw-Hill/Irwin*, 2008 (ISBN 978-0-07-353062-8).
3. Morris G., “Finance Director’s Handbook”, 5th edition, *Elsevier*, 2009 (ISBN-13: 978-0-7506-8701-0)
4. Vernimmen P., Quiry P., Dallochio M., Le Fur Y., and Salvi A., “Corporate Finance Theory and Practice”, 2nd edition, *John Wiley & Sons*, 2009 (ISBN 978-0-470-72192-6).
5. Νιάρχος Ν., “Χρηματοοικονομική Ανάλυση Λογιστικών Καταστάσεων”, 7^η έκδοση, Εκδ. *Σταμούλη Α.Ε.*, Αθήνα, 2004.
6. Higson A., “Corporate Financial Reporting: Theory and Practice”, *SAGE Publications Ltd*, 2003 (ISBN 0-7619-7141-6).
7. Smith J., and Skousen F., “Intermediate Accounting”, 9th edition, *South-Western Publishing Company*, 1987.
8. Atrill P., and McLaney E., “Accounting and Finance for Non-Specialists”, 5th edition, *Prentice Hall*, 2006 (ISBN-13: 978-0-273-70244-3).
9. Vause B., “Guide to Analysing Companies”, 4th edition, *Profile Books LTD*, 2005 (ISBN-13: 978 1 86197 985 8).
10. Fight A., “Credit Risk Management”, *Elsevier Essential Capital Markets*, 2004, (ISBN: 0-7506-5903-3).
11. Bekaert G., and Hodrick R., “International Financial Management”, 2nd edition, *Prentice Hall*, 2012 (ISBN-13: 978-0-13-216276-0).
12. Καφούσης Γ., “Ανάλυση και Κριτική Διερεύνηση Χρηματοοικονομικών Καταστάσεων”, Αθήνα 1991.
13. Glautier M.W.E., and Underdown B., “Accounting Theory and Practice”, 3rd edition, *Pitman Publishing LTD*, 1985.
14. Kennedy R., and McMullen S., “Financial Statements, Forms, Analysis and Interpretation”, 5th edition, *Richard D. Irwin Inc.*, 1968.
15. Titman S., Martin J., and Keown A., “Financial Management: Principles and Applications”, 11th edition, *Prentice Hall*, 2011 (ISBN-13: 9780132340359).

16. Dyson J., “Accounting for Non-Accounting Students”, 8th edition, Financial Times, *Prentice Hall*, 2010 (ISBN: 978-0-273-72297-7).
17. Lasher W., “Practical Financial Management”, 5th edition, *Thomson South-Western*, 2008 (ISBN-13: 978-0-324-42262-7).

ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

3

3.1 Εισαγωγή

Τις τελευταίες δεκαετίες, οι ερευνητές σε θέματα Τεχνητών Νευρωνικών Δικτύων έχουν επιτύχει σημαντική πρόοδο και ανακαλύψεις σε όλες τα πεδία έρευνας αναφορικά με τα νευρωνικά δίκτυα, συμπεριλαμβανομένων των θεωρητικών προσεγγίσεων, των αρχιτεκτονικών, της μοντελοποίησης και προσομοίωσης, των εμπειρικών μελετών, καθώς και σε ένα ευρύ φάσμα εφαρμογών διαφορετικών τομέων. Οι πρόσφατες εξελίξεις της επιστήμης και της τεχνολογίας, συμπεριλαμβανομένων των νευροεπιστημών, της επιστήμης των υπολογιστών, της γνωσιακής επιστήμης, της νανο-τεχνολογίας και του σχεδιασμού της επιστήμης των μηχανικών έδωσαν νέες σημαντικές αντιλήψεις και τεχνολογικές λύσεις για την ώθηση της έρευνας των νευρωνικών δικτύων προς την ανάπτυξη πολύπλοκων, μεγάλης κλίμακας, διασυνδεδεμένων ευφυών συστημάτων που συμπεριφέρονται παρόμοια με τον ανθρώπινο εγκέφαλο.

Τα Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks - ANNs) ή ΤΝΔ εν συντομία, προσελκύουν το έντονο ενδιαφέρον πολλών ερευνητών τις τελευταίες δεκαετίες. Το πλήθος των εφαρμογών στις οποίες έχουν χρησιμοποιηθεί τα ΤΝΔ αποδεικνύει την πολύτιμη συνεισφορά τους και ενδυναμώνει την περαιτέρω έρευνα στην περιοχή αυτή [1, 2, 3].

Τα ΤΝΔ είναι παράλληλα καταναμημένα υπολογιστικά συστήματα αποτελούμενα από απλές προσαρμόσιμες μονάδες επεξεργασίας, ονομαζόμενες νευρώνες πλήρως διασυνδεδεμένες μεταξύ τους, και έχουν την ικανότητα να ανταποκρίνονται σε ερεθίσματα που δέχονται στην είσοδό τους και να μαθαίνουν να προσαρμόζονται στο περιβάλλον τους [3-7]. Η δομή και η λειτουργία των ΤΝΔ είναι εμπνευσμένα από το βιολογικό νευρικό σύστημα των ζώων οργανισμών και ιδιαίτερα του ανθρώπου, τόσο από λειτουργικής όσο και από δομικής πλευράς.

Τα νευρωνικά δίκτυα αποκαλούνται επίσης διασυνδεδεμένα συστήματα, παράλληλα καταναμημένα συστήματα ή προσαρμόσιμα συστήματα λόγω του παράλληλου τρόπου λειτουργίας των μονάδων επεξεργασίας. Στα ΤΝΔ δεν υπάρχει κεντρικός έλεγχος με την κλασική μορφή, καθώς όλες οι διασυνδεδεμένες μονάδες επεξεργασίας τροποποιούνται ή “προσαρμόζονται” ταυτόχρονα με τη ροή της πληροφορίας και τους κανόνες που επικρατούν.

Η υπολογιστική ισχύς ενός ΤΝΔ οφείλεται στη συμπαγή παράλληλη καταναμημένη δομή του και στην ικανότητά του να μαθαίνει και εν συνεχεία να γενικεύει, παράγοντας λογικά αποτελέσματα για δεδομένα που δεν είχε αντιμετωπίσει κατά τη διαδικασία μάθησης.

Τα σημαντικότερα χαρακτηριστικά των τεχνητών νευρωνικών δικτύων είναι:

1. **Μη-γραμμικότητα.** Ένα νευρωνικό δίκτυο έχει μη-γραμμική συμπεριφορά καθώς αποτελείται από ένα σύνολο νευρώνων που είναι μη-γραμμικές μονάδες επεξεργασίας δεδομένων. Η ιδιότητα αυτή επιτρέπει στα ΤΝΔ να υπολογίζουν μη-γραμμικές συναρτήσεις και να εκτελούν πολύπλοκους μετασχηματισμούς στα δεδομένα τους. Το χαρακτηριστικό της μη-γραμμικότητας είναι πολύ σημαντικό δεδομένου ότι ο φυσικός μηχανισμός παραγωγής των σημάτων εισόδου σε ένα νευρωνικό δίκτυο είναι μη-γραμμικός.

2. **Προσαρμοστικότητα.** Τα ΤΝΔ συσχετίζουν δεδομένα εισόδου με τα αντίστοιχα επιθυμητά αποτελέσματα, προσαρμόζοντας τις παραμέτρους τους, στις αλλαγές του περιβάλλοντός τους.

3. **Γενίκευση.** Τα ΤΝΔ δίκτυα έχουν την ικανότητα να ανακαλύπτουν θεμελιώδεις σχέσεις και να χρησιμοποιούν αποθηκευμένη εμπειρική γνώση [4-7], ώστε να μπορούν να γενικεύουν όταν δεχθούν στις εισόδους τους νέα μη-γνώριμα δεδομένα.

4. **Ανοχή σε σφάλματα ή Ευρωστία.** Τα νευρωνικά δίκτυα παρουσιάζουν ανοχή τόσο σε φυσικές βλάβες όσο και στην ύπαρξη θορύβου στα δεδομένα. Αυτό συνεπάγεται την εξασφάλιση της λειτουργίας που επιτελούν τα ΤΝΔ, ακόμη κι αν απαιτηθεί μείωση της απόδοσής τους.

5. **Υλοποίηση σε VLSI.** Η συμπαγής παράλληλη φύση των νευρωνικών δικτύων, καθιστά δυνατή την υλοποίησή τους σε συστήματα τεχνολογίας VLSI (Very Large Scale Integration), έτσι ώστε τα ΤΝΔ να μπορούν να χρησιμοποιηθούν σε εφαρμογές πραγματικού χρόνου.

6. **Αναλογία με Νευροβιολογία.** Κατά τη σχεδίαση των τεχνητών νευρωνικών δικτύων παρατηρείται μια αναλογία με τη δομή του εγκεφάλου. Οι νευροβιολόγοι βλέπουν τα ΤΝΔ σαν αντικείμενο έρευνας για την εξήγηση νευροβιολογικών φαινομένων. Ομοίως, οι μηχανικοί βλέπουν στη νευροβιολογία νέες ιδέες για την επίλυση πολύπλοκων προβλημάτων.

Η αρχιτεκτονική των ΤΝΔ είναι πολύ διαφορετική από αυτή των κλασικών υπολογιστών που περιέχουν έναν επεξεργαστή. Τα νευρωνικά δίκτυα έχουν την ικανότητα για παράλληλη επεξεργασία, όταν υλοποιούνται σε παράλληλους ψηφιακούς υπολογιστές ή προσαρμοσμένο υλικό, σε αντίθεση με την αρχιτεκτονική von Neumann, όπου επικρατεί η σειριακή επεξεργασία [8-10].

Τα νευρωνικά δίκτυα αποτελούν μια διαφορετική μέθοδο για την ανάλυση δεδομένων και την αναγνώριση προτύπων από τις συμβατικές μεθόδους. Επιπρόσθετα, επεξεργάζονται τις πληροφορίες με τρόπο παρόμοιο με αυτόν του ανθρώπινου εγκεφάλου. Καθώς τα ΤΝΔ “μαθαίνουν μέσω παραδειγμάτων”, αποτελούν πολύ ελκυστική τεχνική σε εφαρμογές που παρατηρείται ελάχιστη ή ανολοκλήρωτη κατανόηση του προς επίλυση προβλήματος, εντούτοις, υπάρχουν διαθέσιμα δεδομένα για την προσαρμογή τους.

Τα νευρωνικά δίκτυα και οι συμβατικοί υπολογιστές αρχιτεκτονικής von Neumann πραγματοποιούν πολύ διαφορετικές εργασίες και έχουν διαφορετικές ιδιότητες. Τα ολοκληρωμένα κυκλώματα των υπολογιστών μπορούν να εκτελούν μία εντολή σε χρόνο ενός nanosecond, ενώ οι νευρώνες είναι εκατομμύρια φορές πιο αργοί. Όμως, ο εγκέφαλος, αντισταθμίζει αυτό το μειονέκτημά του, διότι όλοι οι νευρώνες και οι συνάψεις είναι ταυτόχρονα ενεργοί, ενώ οι περισσότεροι σημερινοί υπολογιστές έχουν μόνο μία ή, στην καλύτερη περίπτωση, μερικές κεντρικές μονάδες επεξεργασίας. Συνεπώς, παρόλο που ένας υπολογιστής είναι ένα εκατομμύριο φορές γρηγορότερος σε καθαρή ταχύτητα μεταγωγής, ο εγκέφαλος είναι τελικά 100.000 φορές γρηγορότερος στη δουλειά που κάνει. Ο Πίνακας 3.1 συνοψίζει τα χαρακτηριστικά μεταξύ των τεχνητών νευρωνικών δικτύων και της αρχιτεκτονικής von Neumann.

Τα νευρωνικά δίκτυα και οι συμβατικοί υπολογιστές λειτουργούν συμπληρωματικά. Υπάρχουν διεργασίες καταλληλότερες για αλγοριθμική προσέγγιση, όπως οι αριθμητικές εφαρμογές, και διεργασίες που τα νευρωνικά δίκτυα μπορούν να τις διεκπεραιώσουν αποτελεσματικότερα. Επιπρόσθετα, ένας μεγάλος αριθμός εφαρμογών απαιτεί συστήματα που χρησιμοποιούν ένα συνδυασμό των δυο προσεγγίσεων, με στόχο την επίτευξη μέγιστης αποδοτικότητας.

<i>Νευρωνικά Δίκτυα</i>	<i>Αρχιτεκτονική von Neumann</i>
Εργάζονται με σύγχρονο τρόπο λειτουργίας	Εργάζεται με ασύγχρονο τρόπο λειτουργίας
Παράλληλη επεξεργασία	Σειριακή επεξεργασία
Εκπαιδούνται με παραδείγματα αλλάζοντας τα βάρη των συνδέσεών τους	Προγραμματίζεται με εντολές λογικού χαρακτήρα (if – then)
Η μνήμη, τα δίκτυα και οι μονάδες λειτουργίας συνυπάρχουν	Η μνήμη και επεξεργασία πληροφορίας χωρίζονται
Ανοχή στα σφάλματα	Καμία ανοχή στα σφάλματα
Αυτό-οργάνωση κατά τη διαδικασία της εκπαίδευσης	Εξαρτάται εξ ολοκλήρου από το προσφερόμενο λογισμικό
Η πληροφορία αποθηκεύεται στα βάρη των συνδέσεων	Η πληροφορία αποθηκεύεται σε συγκεκριμένες διευθύνσεις μνήμης
Χρόνος ενός κύκλου: της τάξης του msec	Χρόνος ενός κύκλου: της τάξης του nsec

Πίνακας 3.1. Σύγκριση αρχιτεκτονικής von Neumann και τεχνητών νευρωνικών δικτύων.

3.2 Ιστορική Αναδρομή

Η μελέτη της λειτουργίας του ανθρώπινου εγκεφάλου έχει τις ρίζες της στην αρχαιότητα. Σε έναν αιγυπτιακό πάπυρο που χρονολογείται περίπου στο 3.000 π.Χ. γίνονταν αναφορές για ασθενείς με τραύματα στο κρανίο και τον εγκέφαλο καθώς επίσης και κάποιες προσπάθειες εντοπισμού ορισμένων λειτουργιών στον φλοιό του εγκεφάλου. Όλη η ιστορική αναδρομή των Τεχνητών Νευρωνικών Δικτύων παρουσιάζεται αναλυτικά στα [11-13]. Περιληπτικά παρακάτω αναφέρονται οι κυριότερες ιστορικές στιγμές στην πορεία εξέλιξης των ΤΝΔ.

Η αρχή των νευροεπιστημών θεωρείται η 18^η Μαΐου 1861, όπου ο Γάλλος γιατρός, ανατόμος και ανθρωπολόγος Pierre Paul Broca είπε: “μιλάμε με το αριστερό ημισφαίριο” [14].

Σημαντικό έργο στο πεδίο των νευρωνικών δικτύων έλαβε χώρα στα τέλη του 19^{ου} και στις αρχές του 20^{ου} αιώνα. Το έργο αυτό κυρίως αποτελούνταν από δουλειά που βασιζόταν πάνω σε διαφορετικές αρχές της φυσικής, ψυχολογίας και νευροφυσιολογίας από επιστήμονες όπως οι Hermann von Helmholtz, Ernst Mach και Ivan Pavlov.

Η σύγχρονη περίοδος των νευρωνικών δικτύων [11, 13] άρχισε με την πρωτοποριακή δουλειά του νευροφυσιολόγου Warren McCulloch και του μαθηματικού Walter Pitts το 1943, που δημοσίευσαν μια εργασία για τον τρόπο λειτουργίας των νευρώνων [15]. Η επόμενη μεγάλη ανάπτυξη πάνω στα νευρωνικά δίκτυα, ήρθε από τον Donald Hebb το 1949, με την έκδοση του βιβλίου “The Organization of Behavior”, στο οποίο παρουσιάζεται ο κανόνας μάθησης του Hebb [16].

Καθώς, οι υπολογιστές εξελίσσονταν, στη δεκαετία του 1950 ήταν εφικτή η προσομοίωση ενός υποθετικού νευρωνικού δικτύου. Τα πρώτα βήματα προς αυτή την κατεύθυνση έγιναν από τον Nathaniel Rochester που εργαζόταν στα ερευνητικά εργαστήρια της IBM. Δυστυχώς, για τον ίδιο, η πρώτη του αυτή προσπάθεια απέτυχε, αλλά τα μετέπειτα εγχειρήματα ήταν επιτυχή. Ήταν η περίοδος που άνθιζαν οι κλασικοί υπολογιστές, παραμελώντας την έρευνα για τα νευρωνικά δίκτυα [17].

Το 1954, ο Minsky έγραψε τη διδακτορική του διατριβή με τίτλο “Theory of Neural-Analog Reinforcement Systems and Its Application to the Brain-Model Problem” [18] και το 1961 ο ίδιος έγραψε μια εργασία με τίτλο “Steps Toward Artificial Intelligence” [19].

Το 1956, το “Dartmouth Summer Research Project on Artificial Intelligence” έδωσε ώθηση στην τεχνητή νοημοσύνη και στα νευρωνικά δίκτυα [20]. Η εργασία των Rochester, Holland, Habit και Duda [21] είναι ίσως η πρώτη προσπάθεια για εξομοίωση με τη χρήση υπολογιστή ώστε να ελεγχθεί μια καλά σχηματισμένη νευρωνική θεωρία βασισμένη στον κανόνα μάθησης του Hebb. Τον ίδιο χρόνο ο Uttley παρουσίασε την αποκαλούμενη “διαρρέουσα ολοκλήρωση” ή “νευρώνα φωτιάς” που αργότερα παρουσιάστηκε κι από τον Caianiello [13].

Η πρώτη πρακτική εφαρμογή των τεχνητών νευρωνικών δικτύων παρουσιάστηκε στα τέλη της δεκαετίας του '50 με την εφεύρεση του perceptron δικτύου και του συσχετισμένου κανόνα

μάθησης από τον Frank Rosenblatt [22], 15 χρόνια μετά από τη δημοσίευση της εργασίας των McCulloch και Pitts. Ο Rosenblatt και οι συνάδελφοί του κατασκεύασαν ένα perceptron δίκτυο και παρουσίασαν την ικανότητά του να πραγματοποιεί ταξινόμηση προτύπων. Αυτή η πρώιμη επιτυχία δημιούργησε ένα έντονο ενδιαφέρον για έρευνα στο πεδίο των νευρωνικών δικτύων καθώς άρχισε να συζητείται η ιδέα ότι πιθανώς τα νευρωνικά δίκτυα είναι η τεχνική που μπορεί να λύσει όλα τα προβλήματα που μέχρι τότε παρέμεναν άλυτα. Δυστυχώς, αργότερα αποκαλύφθηκε ότι το βασικό δίκτυο perceptron μπορούσε να επιλύσει μόνο μια περιορισμένη κατηγορία προβλημάτων.

Το 1960, οι Bernard Widrow και Marcian Hoff από το πανεπιστήμιο του Stanford, παρουσίασαν τον αλγόριθμο ελάχιστου μέσου τετραγώνου (least mean square – LMS), ο οποίος χρησιμοποιείται ακόμη και σήμερα για την εκπαίδευση των νευρωνικών δικτύων [23].

Στα τέλη της δεκαετίας του '60, το ενδιαφέρον για τα νευρωνικά δίκτυα εξασθένησε, παρόλο που επικρατούσε η άποψη ότι δεν υφίσταντο περιορισμοί αναφορικά με τις δυνατότητές τους και ότι μπορούσαν να δώσουν λύσεις σε όλα τα υπάρχοντα προβλήματα του πραγματικού κόσμου. Την περίοδο αυτή εκδίδεται το βιβλίο των Marvin Minsky και Seymour Papert [24] με τίτλο “Perceptrons”, που αποδείκνυε μαθηματικά ότι υπάρχουν όρια στις υπολογιστικές ικανότητες των αισθητηρίων (perceptrons).

Πολλοί άνθρωποι, επηρεάστηκαν από τους Minsky και Papert, με αποτέλεσμα να πιστέψουν ότι περεταίρω έρευνα στα νευρωνικά δίκτυα οδηγούσε σε αδιέξοδο. Αυτό, σε συνδυασμό με το γεγονός ότι δεν υπήρχαν ισχυροί ψηφιακοί υπολογιστές για την πραγματοποίηση πειραμάτων, έγιναν αιτία ώστε πολλοί ερευνητές να εγκαταλείψουν το πεδίο των νευρωνικών δικτύων. Για μια δεκαετία η έρευνα στα νευρωνικά δίκτυα είχε σε μεγάλο βαθμό ανασταλεί.

Εντούτοις, σημαντική έρευνα εξακολούθησε να γίνεται κατά τη διάρκεια της δεκαετίας του '70. Συγκεκριμένα, το 1972, ο Teuvo Kohonen [25] και ο James Anderson [26], ανεξάρτητα ο ένας από τον άλλο, ανέπτυξαν νέα παρόμοια νευρωνικά δίκτυα, που μπορούσαν να δράσουν σαν μνήμες. Το πρώτο πολυστρωματικό (multi-layer) δίκτυο αναπτύχθηκε το 1975 και ήταν ένα δίκτυο μη επιβλεπόμενης μάθησης. Επίσης, στο ίδιο διάστημα, ο Stephen Grossberg [27] υπήρξε πολύ δραστήριος στην έρευνα των αυτό-οργανούμενων δικτύων (self-organizing networks).

Ανάκαμψη του ενδιαφέροντος για το πεδίο των νευρωνικών δικτύων παρατηρείται στις αρχές της δεκαετίας του '80, μετά από την παρουσίαση νέων ιδεών και την αύξηση των δυνατοτήτων των υπολογιστών. Ο αριθμός των επιστημόνων, των συνεδρίων και των περιοδικών και τα ποσά που δαπανήθηκαν σχετικά με τα νευρωνικά δίκτυα, καταδεικνύουν την ανοδική πορεία του συγκεκριμένου ερευνητικού τομέα [13, 28]. Τη δεκαετία του '80 έγιναν πολλές εργασίες πάνω στη θεωρία αλλά και στο σχεδιασμό των νευρωνικών δικτύων.

Το 1982 ο John Hopfield βασιζόμενος στη στατιστική μηχανική, χρησιμοποίησε την ιδέα της συνάρτησης ενέργειας για να φτιάξει ένα νέο τρόπο κατανόησης του υπολογισμού που γίνεται

από τα δίκτυα με ανάδραση (recurrent networks) [29]. Η προσέγγιση του Hopfield δεν ήταν απλά η μοντελοποίηση του εγκεφάλου, αλλά η δημιουργία χρήσιμων συσκευών. Η κατηγορία νευρωνικών δικτύων με ανατροφοδότηση έτυχε ιδιαίτερης προσοχής κατά τη δεκαετία του '80, και με το χρόνο έγιναν γνωστά σαν δίκτυα Hopfield. Μια ακόμα σημαντική ανάπτυξη το 1982 ήταν η έκδοση της εργασίας του Kohonen πάνω στους αυτό-οργανώμενους χάρτες (Self-Organizing Feature Maps – SOFMs), χρησιμοποιώντας μιας ή δύο διαστάσεων δικτυωτές δομές.

Επίσης, το 1982 πραγματοποιήθηκε ένα συνέδριο, με τη συνεργασία Ηνωμένων Πολιτειών και Ιαπωνίας, για Συνεργατικά / Ανταγωνιστικά Νευρωνικά Δίκτυα (Cooperative / Competitive Neural Networks) στο Κιότο της Ιαπωνίας [11, 13]. Η Ιαπωνία ανακοίνωσε μια νέα, πέμπτης γενιάς προσπάθεια επί των νευρωνικών δικτύων, ενώ οι εργασίες που παρουσιάστηκαν από την πλευρά των Ηνωμένων Πολιτειών δημιούργησαν μια ανησυχία ότι είχαν μείνει πίσω στο συγκεκριμένο πεδίο. Συνεπώς, υπήρξε επιπρόσθετη χρηματοδότηση και κατ' επέκταση περισσότερη έρευνα στο συγκεκριμένο πεδίο [11].

Το 1986, τα *πολυστρωματικά perceptron δίκτυα (Multi-Layer Perceptrons – MLPs)*, είναι αυτά που κυριαρχούν στο πεδίο των νευρωνικών δικτύων. Όμως, ο αλγόριθμος των Widrow-Hoff δεν μπορούσε να χρησιμοποιηθεί για την εκπαίδευσή τους, με αποτέλεσμα την ανάπτυξη του αλγόριθμου οπισθοδιάδοσης σφάλματος (back-propagation error algorithm) από διαφορετικούς επιστήμονες. Η επίσημη δημοσίευση για τον αλγόριθμο οπισθοδιάδοσης σφάλματος έγινε από τους David Rumelhart και James McClelland [30]. Ο αλγόριθμος αυτός αποτελούσε την απάντηση στις επικρίσεις που έκαναν οι Minsky και Papert τη δεκαετία του '60. Ο συγκεκριμένος αλγόριθμος έγινε πολύ δημοφιλής και έδωσε νέα ώθηση στις εφαρμογές των νευρωνικών δικτύων.

Παρ' όλο που είναι δύσκολο να προβλεφθεί η μελλοντική επιτυχία των νευρωνικών δικτύων, ο μεγάλος αριθμός και η ευρεία περιοχή εφαρμογών αυτής της νέας τεχνολογίας είναι πολύ ενθαρρυντικά. Στην επόμενη ενότητα παρουσιάζονται οι τομείς που μπορούν να βρουν εφαρμογή τα τεχνητά νευρωνικά δίκτυα.

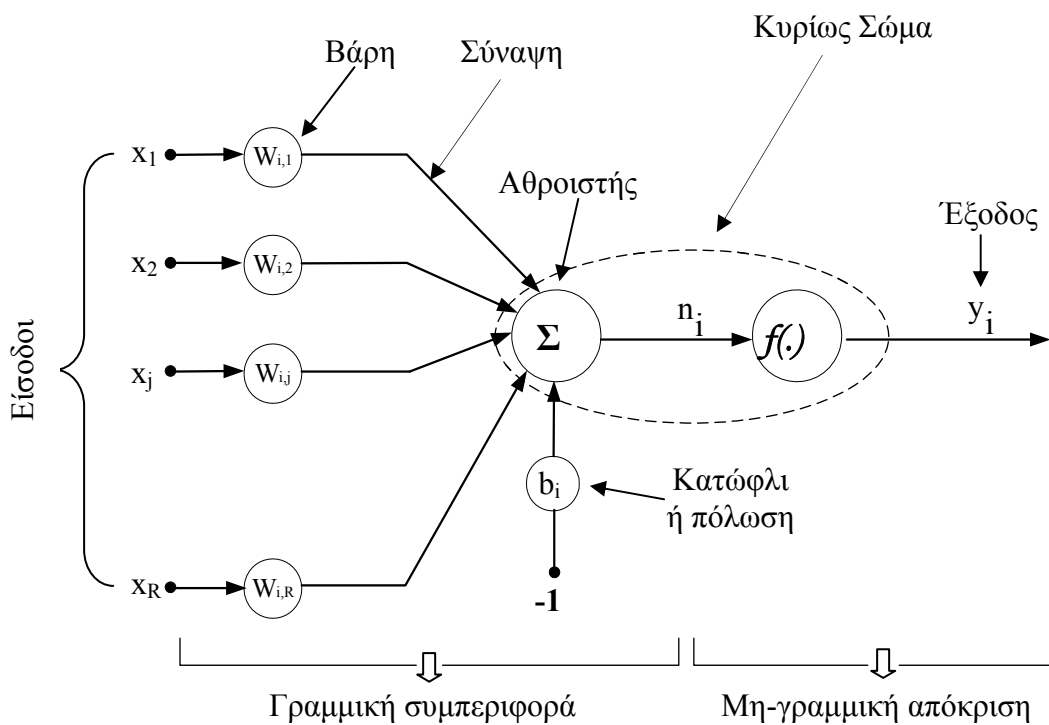
3.3. Μοντέλο Τεχνητού Νευρώνα

Η θεμελιώδης μονάδα για τη λειτουργία και την επεξεργασία της πληροφορίας ενός νευρωνικού δικτύου είναι ο *νευρώνας*. Κάθε νευρώνας δέχεται ένα πλήθος δεδομένων σαν εισόδους μέσω ενός συνόλου συνδέσεων που ονομάζονται *συνάψεις*. Το μοντέλο ενός τεχνητού νευρώνα απεικονίζεται στο Σχήμα 3.1, όπου διακρίνονται οι συνάψεις, το κυρίως σώμα και η έξοδος του νευρώνα.

Οι *συνάψεις* αποτελούν το μέσο για τη λήψη των πληροφοριών που θα οδηγηθούν στο κυρίως σώμα του νευρώνα για επεξεργασία. Τα δεδομένα μπορεί να προέρχονται είτε από εξωτερικές πηγές είτε από άλλους νευρώνες. Σε κάθε σύναψη αντιστοιχεί μια τιμή που αποτελεί

το συντελεστή βαρύτητας του αντίστοιχου λαμβανόμενου σήματος και ονομάζεται **συναπτικό βάρος ή βάρος**.

Το **κυρίως σώμα** του νευρώνα περιλαμβάνει έναν **κόμβο άθροισης (summation junction), Σ** , και μια **συνάρτηση ενεργοποίησης, f** . Ο αθροιστής προσθέτει αλγεβρικά τα σταθμισμένα λαμβανόμενα σήματα που δέχθηκε ο νευρώνας στις εισόδους του. Η συνάρτηση ενεργοποίησης εκτελεί μια μαθηματική πράξη στο προκύπτον σήμα (n_i), ενώ παράλληλα διαμορφώνει το σήμα εξόδου του νευρώνα (y_i).



Σχήμα 3.1. Μοντέλο τεχνητού νευρώνα.

Στο Σχήμα 3.1 διακρίνονται δυο βασικά τμήματα του νευρώνα, ένα γραμμικής αντιστοίχισης (linear map) κι ένα μη-γραμμικής απόκρισης (non-linear map). Το πρώτο τμήμα υπολογίζει τη **γραμμική συνδυαστική έξοδο του νευρώνα** (n_i) αθροίζοντας αλγεβρικά τα γινόμενα κάθε σήματος εισόδου (x_j) με το αντίστοιχο συναπτικό βάρος ($w_{i,j}$) και μια ποσότητα αντιστάθμισης, που ονομάζεται **κατώφλι** ή **πόλωση** b_i και επιδρά στην ενεργοποίηση της εξόδου του νευρώνα. Η τιμή του σήματος δίνεται από τη σχέση:

$$n_i = \sum_{j=1}^R w_{i,j} x_j + b_i \quad (3.1)$$

Το τμήμα της μη-γραμμικής απόκρισης μετασχηματίζει το σήμα n_i ώστε να προκύψει η έξοδος του νευρώνα y_i , σύμφωνα με την ακόλουθη σχέση:

$$y_i = f \left(\sum_{j=1}^R w_{i,j} x_j + b_i \right) \quad (3.2)$$

3.4 Συναρτήσεις Μεταφοράς

Η έρευνα για τα νευρωνικά δίκτυα δίνει ιδιαίτερη έμφαση στους κανόνες εκμάθησης, που προσδιορίζουν τον τρόπο εκπαίδευσης των νευρωνικών δικτύων και στις αρχιτεκτονικές τους, που είναι ο τρόπος διάταξης των νευρώνων και οι μεταξύ τους συνδέσεις, αδιαφορώντας για τον καθοριστικό ρόλο των *συναρτήσεων μεταφοράς* [14, 15].

Οι συναρτήσεις μεταφοράς διακρίνονται σε *συναρτήσεις ενεργοποίησης* και *συναρτήσεις εξόδου*. Η *συνάρτηση ενεργοποίησης* προσδιορίζει το σήμα που λαμβάνει ο νευρώνας από τις εισόδους του. Η τιμή της συνάρτησης ενεργοποίησης είναι συνήθως βαθμωτό μέγεθος και τα ορίσματά της διανύσματα. Η *συνάρτηση εξόδου $o(I)$* καθορίζει την επεξεργασία του σήματος από το νευρώνα, λειτουργώντας με βαθμωτές ενεργοποιήσεις και επιστρέφοντας βαθμωτές τιμές. Οι δυο αυτοί τύποι συναρτήσεων καθορίζουν τον τρόπο επεξεργασίας των σημάτων και διαμορφώνουν τις τιμές των σημάτων εξόδου των νευρώνων. Ο συνδυασμός της συνάρτησης *ενεργοποίησης* και της συνάρτησης *εξόδου* καλείται *συνάρτηση μεταφοράς $o(I(x))$* [15].

Οι συναρτήσεις ενεργοποίησης και εξόδου διαφέρουν μεταξύ των επιπέδων των νευρωνικών δικτύων [15]. Ο Πίνακας 3.2 συνοψίζει όλες τις συχνότερα χρησιμοποιούμενες συναρτήσεις εξόδου. Η 2^η στήλη του πίνακα περιγράφει τη συνάρτηση. Η 3^η στήλη δίνει τη μαθηματική έκφραση των συναρτήσεων εξόδου [14, 31].

Επισημαίνεται ότι μεγάλο μέρος της βιβλιογραφίας που αναφέρεται στα νευρωνικά δίκτυα, ταυτίζει τους όρους *συνάρτηση ενεργοποίησης*, *συνάρτηση εξόδου* και *συνάρτηση μεταφοράς* χωρίς καμία διάκριση μεταξύ τους.

<i>ΑΑ</i>	<i>Ονομασία Συνάρτησης</i>	<i>Μαθηματική Εξίσωση</i>
1	Ανταγωνιστική Συνάρτηση	
2	Hard Limit	$f(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$
3	Συμμετρική Hard Limit	$f(x) = \begin{cases} -1, & x < 0 \\ 1, & x \geq 0 \end{cases}$
4	Log sigmoid	$f(x) = \frac{1}{1 + e^{-x}}$
5	Αντίστροφη	$f(x) = \frac{1}{x}$
6	Θετικά Γραμμική	$f(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases}$
7	Γραμμική	$f(x) = x$
8	Ακτινικής Βάσης	$f(x) = e^{-x^2}$
9	Κορεσμένη Γραμμική	$f(x) = \begin{cases} 0, & x \leq 0 \\ x, & 0 \leq x \leq 1 \\ 1, & x \geq 1 \end{cases}$
10	Συμμετρική Κορεσμένη Γραμμική	$f(x) = \begin{cases} -1, & x \leq -1 \\ x, & 0 \leq x \leq 1 \\ 1, & x \geq 1 \end{cases}$
11	Soft Max	$f(x_i) = \frac{e^{x_i}}{\sum_{j=1}^n e^{x_j}}$
12	Hyperbolic Tangent Sigmoid	$f(x) = \frac{2}{1 + e^{-2x}} - 1$
13	Τριγωνικής Βάσης	$f(x) = \begin{cases} 0, & x < -1 \\ 1 - x , & -1 \leq x \leq 1 \\ 0, & x > 1 \end{cases}$

Πίνακας 3.2. Συναρτήσεις εξόδου και οι μαθηματικές τους εκφράσεις [32].

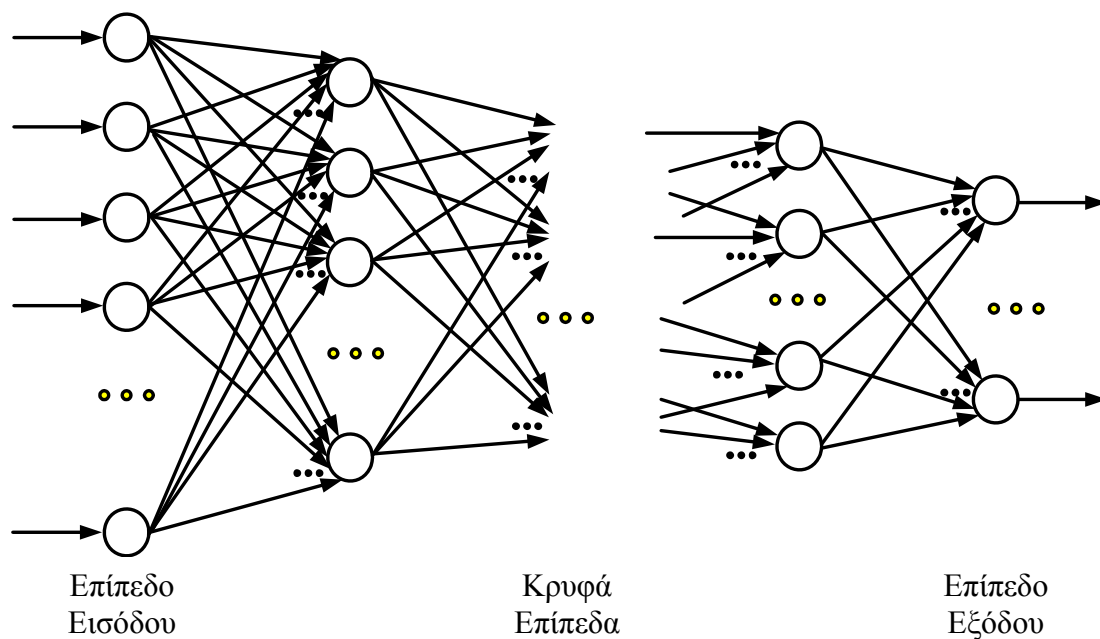
3.5 Τοπολογίες Νευρωνικών Δικτύων

Η τοπολογία ενός νευρωνικού δικτύου εκφράζεται με τον τρόπο που δομούνται οι νευρώνες του. Κάθε τοπολογία έχει τα δικά της χαρακτηριστικά, τα οποία παρουσιάζονται παρακάτω.

3.5.1 Επίπεδα Νευρωνικού Δικτύου

Οι νευρώνες ενός ΤΝΔ κατανέμονται σε *επίπεδα* ή *στρώματα (layers)*. Ένα νευρωνικό δίκτυο αποτελείται από τρία επίπεδα: το *επίπεδο εισόδου (input layer)*, το *κρυφό επίπεδο (hidden layer)* και το *επίπεδο εξόδου (output layer)*. Το Σχήμα 3.2 παρουσιάζει τα επίπεδα ενός νευρωνικού δικτύου.

Το *επίπεδο εισόδου* δέχεται τα δεδομένα από το εξωτερικό περιβάλλον, που θα τροφοδοτηθούν στο δίκτυο για την επεξεργασία τους. Ο αριθμός των νευρώνων στο επίπεδο εισόδου ισούται με τον αριθμό των παραμέτρων του προβλήματος [4, 6].



Σχήμα 3.2. Τα επίπεδα ενός νευρωνικού δικτύου.

Το *κρυφό επίπεδο* είναι αυτό που παρεμβάλλεται μεταξύ του επιπέδου εισόδου και του επιπέδου εξόδου και δεν έχει άμεση επαφή με το περιβάλλον. Ένα νευρωνικό δίκτυο μπορεί να έχει περισσότερα από ένα κρυφά επίπεδα. Το πλήθος των κρυφών επιπέδων και των νευρώνων κάθε επιπέδου είναι ανάλογο της ικανότητας του νευρωνικού δικτύου να προσεγγίζει πολύπλοκες λειτουργίες. Όμως, αυτό δε σημαίνει ότι ένα δίκτυο με πολύπλοκη δομή θα έχει καλύτερη απόδοση, διότι αποκτά ευαισθησία στο θόρυβο, με αποτέλεσμα να μαθαίνει εκτός από την κύρια

λειτουργία του, το θόρυβο που υπάρχει στα δεδομένα εισόδου. Συνεπώς, υπάρχει μια αλληλένδετη σχέση μεταξύ της ικανότητας αναπαράστασης των δεδομένων ενός δικτύου και του θορύβου που το τελευταίο θα ενσωματώσει.

Το **επίπεδο εξόδου** είναι αυτό που παράγει τα διανύσματα εξόδου του νευρωνικού δικτύου. Το επίπεδο εξόδου, όπως και το επίπεδο εισόδου είναι αυτά που έρχονται σε επαφή με το περιβάλλον, για την εμφάνιση των αποτελεσμάτων που προέκυψαν από την επεξεργασία των ληφθέντων δεδομένων. Οι ζητούμενες πληροφορίες μιας εφαρμογής είναι αυτές που καθορίζουν τον αριθμό των νευρώνων στο επίπεδο εξόδου [4, 6, 8].

Επιπρόσθετα, ανάλογα με το πλήθος των επιπέδων τα νευρωνικά δίκτυα διακρίνονται σε **ενός στρώματος (single-layer)** και **πολλαπλών στρωμάτων (multilayer) ή πολυστρωματικά ή πολυ-επίπεδα**.

3.5.2 Στατικά και Δυναμικά Νευρωνικά Δίκτυα

Τα νευρωνικά δίκτυα διακρίνονται σε **στατικά (static)** και **δυναμικά (dynamic ή recurrent)** βάση της δυνατότητας ανατροφοδότησης των εξόδων του νευρωνικού δικτύου.

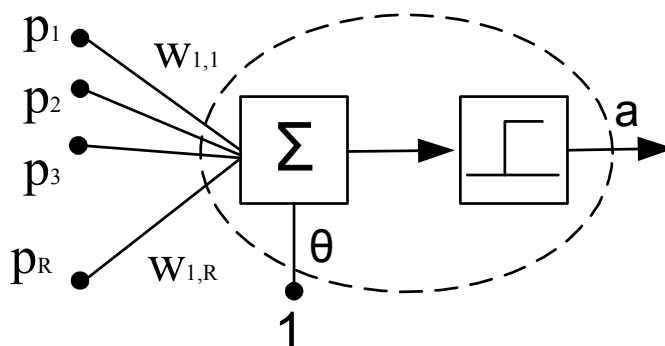
Στα **στατικά** νευρωνικά δίκτυα τα δεδομένα κινούνται μόνο προς μια κατεύθυνση, προς τα εμπρός (forward) από το επίπεδο εισόδου μέσω των κρυφών επιπέδων προς το επίπεδο εξόδου [4, 32-36]. Λόγω της εμπρόσθιας μετάδοσης των δεδομένων, τα στατικά δίκτυα αποκαλούνται **δίκτυα εμπρόσθιας τροφοδότησης (feedforward neural networks)**. Τα δίκτυα αυτά δεν έχουν ανάδραση, με αποτέλεσμα τα δεδομένα εισόδου να είναι ανεξάρτητα της χρονικής ακολουθίας εισαγωγής τους στο νευρωνικό δίκτυο [32].

Τα **δυναμικά** νευρωνικά δίκτυα (**Recurrent Neural Networks – RNNs**) είναι μοντέλα που τα δεδομένα μετακινούνται και προς τις δυο κατευθύνσεις. Ενώ ένα εμπρόσθιας διάδοσης νευρωνικό δίκτυο διαδίδει τα δεδομένα του προς μια κατεύθυνση, το δυναμικό δίκτυο έχει ανάδραση, με αποτέλεσμα τα δεδομένα εξόδου να τροφοδοτούν την είσοδό του. Τα δίκτυα αυτά είναι χρήσιμα για την επεξεργασία χρονικά εξαρτημένων σημάτων [32]. Ένα δυναμικό δίκτυο μπορεί να θεωρηθεί σαν στατικό δίκτυο με στοιχεία καθυστέρησης (delay elements) που αποθηκεύουν προηγούμενες τιμές του σήματος εισόδου.

3.5.3 Αισθητήριο

Το **αισθητήριο (perceptron)** ήταν πιθανά η πρώτη επιτυχής προσπάθεια δημιουργίας νευρωνικού δικτύου. Το perceptron είναι ένα νευρωνικό δίκτυο εμπρόσθιας διάδοσης χωρίς κρυφό επίπεδο που χρησιμοποιείται για την ταξινόμηση γραμμικά διαχωριζόμενων προτύπων [4, 32]. Το δίκτυο perceptron έχει ένα επίπεδο εισόδου και ένα επίπεδο εξόδου που οι νευρώνες του έχουν σαν συνάρτηση εξόδου την hard limit.

Ένας νευρώνας perceptron που χρησιμοποιεί σαν συνάρτηση εξόδου την hard limit παρουσιάζεται στο Σχήμα 3.3α. Διακρίνεται η είσοδος του με ένα διάνυσμα R στοιχείων και ο νευρώνας perceptron.



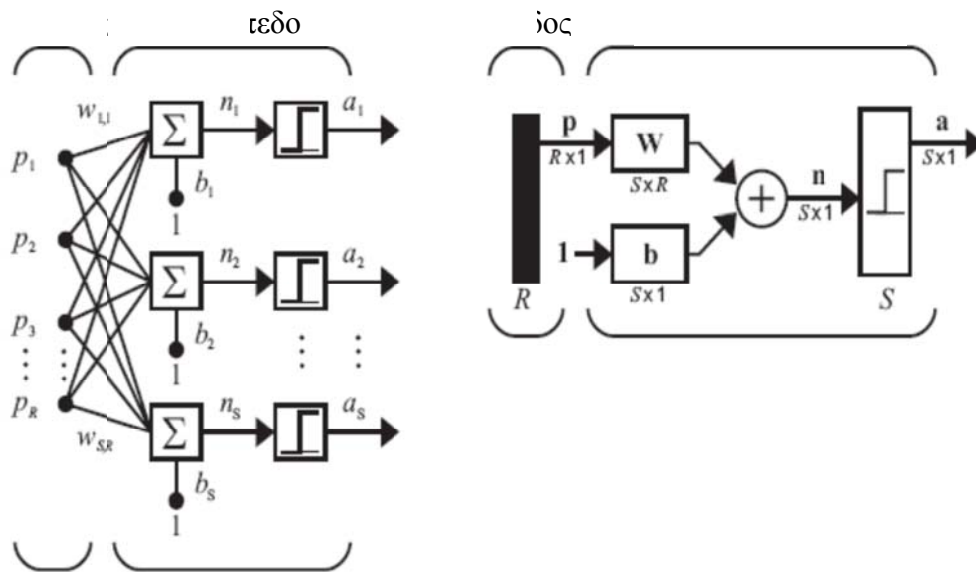
Σχήμα 3.3α. Νευρώνας Perceptron [32].

Σε κάθε στοιχείο εισόδου αντιστοιχεί μια τιμή βάρους, που συμβολίζεται με w_{ij} . Ο δείκτης i δείχνει το νευρώνα στον οποίο καταλήγει το j στοιχείο του διανύσματος εισόδου. Με τον τρόπο αυτό σχηματίζεται ένας πίνακας \mathbf{W} διαστάσεων $S \times R$, όπου το S ισούται με το πλήθος των νευρώνων στο επίπεδο που ακολουθεί και το R με το πλήθος των στοιχείων του διανύσματος εισόδου. Στο Σχήμα 3.3α ο πίνακας \mathbf{W} έχει διαστάσεις $1 \times R$.

Η έξοδος του νευρώνα υπολογίζεται από τη σχέση

$$a = \text{hardlim}(\mathbf{Wp} + \theta)$$

Το Σχήμα 3.3β απεικονίζει ένα perceptron με περισσότερους από ένα νευρώνες. Οι δυο απεικονίσεις είναι ακριβώς ταυτόσημες. Η διαφορά τους έγκειται στον τρόπο αποτύπωσης της απαραίτητης πληροφορίας για το δίκτυο perceptron. Το αριστερό σχήμα δείχνει κάθε κόμβο ξεχωριστά με τις συνδέσεις, τα βάρη, την πόλωση, τη συνάρτηση εξόδου και την έξοδο. Από την άλλη πλευρά, το δεξί σχήμα απεικονίζει το ίδιο ακριβώς δίκτυο, αλλά πιο περιεκτικά. Συγκεκριμένα, αντί να απεικονίσει τους S κόμβους άθροισης και τις S συναρτήσεις εξόδου, δημιουργεί έναν αθροιστή και μια συνάρτηση εξόδου κι αναφέρει το μέγεθός τους. Η δεξιά αναπαράσταση είναι αυτή που χρησιμοποιείται λόγω της ευκολίας στη γραφική αποτύπωση των νευρωνικών δικτύων.



Σχήμα 3.3β. Δίκτυο Perceptron [32].

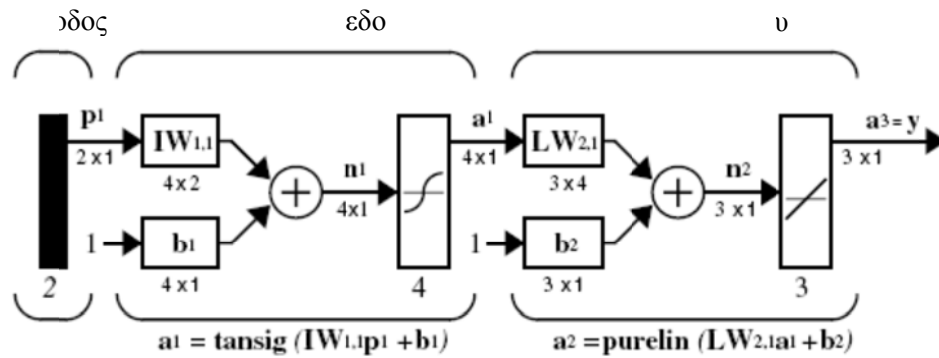
Ο νευρώνας perceptron παράγει ένα (1) αν η είσοδός του είναι μεγαλύτερη ή ίση του μηδενός, διαφορετικά η έξοδός του ισούται με το μηδέν (0). Η συνάρτηση hard limit δίνει τη δυνατότητα στο perceptron να ταξινομεί τα διανύσματα εισόδου χωρίζοντας το χώρο εισόδου (input space) σε δυο περιοχές. Συγκεκριμένα, αν η είσοδος n είναι μικρότερη του μηδενός, η έξοδος θα ισούται με μηδέν, διαφορετικά, αν το n μεγαλύτερο ή ίσο από το μηδέν, η έξοδος του δικτύου θα είναι ένα (1).

3.5.4 Πολλαπλών Στρωμάτων Νευρωνικά Δίκτυα

Τα *πολλαπλών στρωμάτων perceptrons (Multilayer Perceptrons - MLPs)* είναι μια από τις ευρέως χρησιμοποιούμενες τοπολογίες νευρωνικών δικτύων. Τα MLPs αποτελούνται από περισσότερα από ένα στρώματα νευρώνων και χρησιμοποιούν μη-γραμμικά διαφορίσιμες συναρτήσεις εξόδου [7, 22, 32]. Τα MLPs είναι νευρωνικά δίκτυα εμπρόσθιας διάδοσης (feedforward) που μπορούν να χρησιμοποιηθούν σε εφαρμογές που τα διανύσματα εισόδου είναι μη-γραμμικά διαχωρίσιμα. Το Σχήμα 3.4 παρουσιάζει ένα πολυστρωματικό νευρωνικό δίκτυο (MLP).

Ένα δίκτυο MLP έχει συχνά ένα ή περισσότερα κρυφά επίπεδα με σιγμοειδής νευρώνες κι ένα επίπεδο εξόδου με γραμμικούς νευρώνες. Τα MLPs με μη-γραμμικές συναρτήσεις εξόδου έχουν την ικανότητα να μαθαίνουν μη-γραμμικές και γραμμικές συσχετίσεις μεταξύ των διανυσμάτων εισόδου και εξόδου. Το γραμμικό επίπεδο εξόδου επιτρέπει στο δίκτυο να παράγει τιμές που κυμαίνονται από -1 έως 1.

Στην περίπτωση που απαιτείται περιορισμός των εξόδων ενός δικτύου, τότε πρέπει σαν συνάρτηση εξόδου του επιπέδου εξόδου να χρησιμοποιηθεί μια σιγμοειδής συνάρτηση, όπως η log-sigmoid.



Σχήμα 3.4. Αρχιτεκτονική MLP [32].

Η αρχιτεκτονική ενός MLP δεν είναι πλήρως προσδιορισμένη από το πρόβλημα [4, 6, 8]. Ενώ ο αριθμός των νευρώνων για τα επίπεδα εισόδου και εξόδου καθορίζονται από το πρόβλημα, ανάλογα με τις παραμέτρους των διανυσμάτων εισόδου και την επιθυμητή έκβαση, απαιτείται ο προσδιορισμός ενός πλήθους χαρακτηριστικών [7]. Ενδεικτικά χαρακτηριστικά που πρέπει να προσδιορισθούν είναι ο αριθμός των κρυφών επιπέδων και των νευρώνων τους και οι συναρτήσεις εξόδου τους.

Τα MLPs είναι ευαίσθητα στον αριθμό των νευρώνων των κρυφών επιπέδων τους. Συγκριμένα, μικρός αριθμός νευρώνων έχει σαν αποτέλεσμα το νευρωνικό δίκτυο να μην μπορεί να μάθει, ενώ ένας μεγάλος αριθμός νευρώνων οδηγεί στην αδυναμία του δικτύου να γενικεύει, αναγνωρίζοντας μόνο τα δείγματα με τα οποία εκπαιδεύτηκε (overfitting).

3.5.5 Δίκτυα Ακτινικής Βάσης

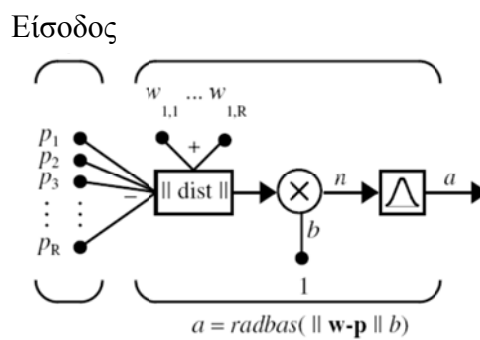
Τα δίκτυα ακτινικής βάσης (*radial basis networks*) έχουν παρόμοια δομή αλλά διαφορετικό μοντέλο νευρώνα από τα MLPs [6, 37]. Ενώ τα δίκτυα ακτινικής βάσης απαιτούν περισσότερους νευρώνες συγκριτικά με τα MLPs, ο χρόνος υλοποίησής τους είναι πολύ λιγότερος από αυτόν που χρειάζεται για τα MLPs. Τα δίκτυα αυτά λειτουργούν καλύτερα όταν υπάρχουν διαθέσιμα πολλά πρότυπα για την εκπαίδευσή τους.

Το Σχήμα 3.5 παρουσιάζει ένα νευρώνα ακτινικής βάσης. Ο υπολογισμός της εισόδου του νευρώνα διαφέρει από εκείνον των MLPs ή του perceptron. Εδώ, η είσοδος του νευρώνα είναι το διάνυσμα της απόστασης μεταξύ των βαρών του και του διανύσματος εισόδου, πολλαπλασιασμένο

με την πόλωση b . Το $\|dist\|$ στο Σχήμα 3.5 δέχεται το διάνυσμα εισόδου \mathbf{p} και την αντίστοιχη γραμμή του πίνακα βαρών, και παράγει το εσωτερικό γινόμενο των δύο.

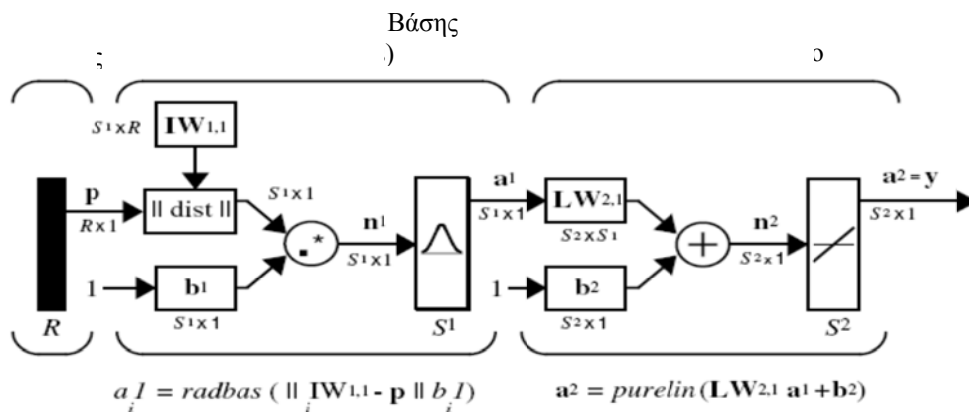
Η συνάρτηση εξόδου που χρησιμοποιείται είναι η radial basis function (RBF) που έχει μέγιστο ίσο με 1 όταν η είσοδος της είναι 0. Καθώς η απόσταση μεταξύ \mathbf{w} και \mathbf{p} μειώνεται, η έξοδος αυξάνεται. Συνεπώς, ένας νευρώνας ακτινικής βάσης λειτουργεί σαν ανιχνευτής (detector, ή φωρατής) που παράγει 1 όταν η είσοδος \mathbf{p} είναι ακριβώς ίδια με το βάρος της.

Η πόλωση b χρησιμοποιείται για την προσαρμογή της ευαισθησίας του ακτινικής βάσης νευρώνα.



Σχήμα 3.5. Νευρώνας Radial Basis δικτύου [32].

Ένα δίκτυο ακτινικής βάσης αποτελείται από τρία επίπεδα. Το επίπεδο εισόδου, ένα κρυφό επίπεδο με ακτινική συνάρτηση βάσης αποτελούμενο από S^1 νευρώνες και ένα γραμμικό επίπεδο εξόδου με S^2 νευρώνες (Σχήμα 3.6).

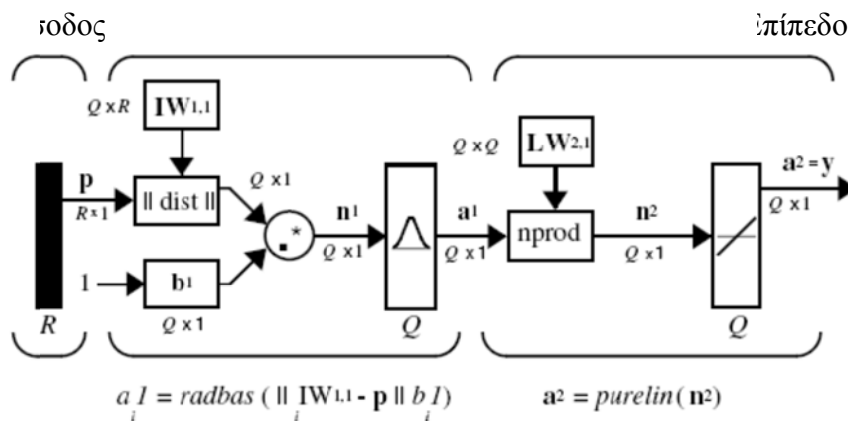


Σχήμα 3.6. Αρχιτεκτονική Radial Basis δικτύου [32].

Η λειτουργία του κρυφού επιπέδου είναι η ακόλουθη. Η σταθμισμένη είσοδος (weighted input) κάθε νευρώνα είναι η απόσταση μεταξύ του διανύσματος εισόδου και του διανύσματος βάρους του, υπολογιζόμενη με την $\|dist\|$. Η είσοδος που οδηγείται στη συνάρτηση εξόδου είναι το στοιχείο-προς-στοιχείο γινόμενο της σταθμισμένης εισόδου με την πόλωση b . Η έξοδος κάθε νευρώνα είναι το αποτέλεσμα που προκύπτει όταν το σήμα εισόδου περάσει από την ακτινική συνάρτηση βάσης. Αν το διάνυσμα βάρους ενός νευρώνα ισούται με το διάνυσμα εισόδου, τότε η σταθμισμένη είσοδος είναι ίση με το 0 και η έξοδός του είναι 1. Αν το διάνυσμα βάρους του νευρώνα έχει μια απόσταση ανοίγματος (distance of spread) από το διάνυσμα εισόδου, τότε η σταθμισμένη είσοδος ισούται με την τιμή ανοίγματος, και η έξοδός του είναι 0.5.

Δυο παραλλαγές των δικτύων ακτινικής βάσης είναι τα **δίκτυα γενικευμένης παλινδρόμησης (Generalized Regression Neural Networks – GRNNs)** και τα **πιθανοκρατικά νευρωνικά δίκτυα (Probabilistic Neural Networks- PNNs)** [5].

Ένα **GRNN** έχει ένα RBF κρυφό επίπεδο και ένα ειδικό γραμμικό επίπεδο εξόδου. Συνήθως, χρησιμοποιείται για προσέγγιση συναρτήσεων. Το GRNN είναι όμοιο με το δίκτυο ακτινικής βάσης, αλλά έχει μια μικρή διαφοροποίηση στο επίπεδο εξόδου. Το Σχήμα 3.7 παρουσιάζει ένα GRNN. Εδώ υπάρχει το nprod, που κανονικοποιεί το εσωτερικό γινόμενο των βαρών. Κάθε στοιχείο είναι το εσωτερικό γινόμενο του $LW^{2,1}$ με το διάνυσμα a^1 , κανονικοποιημένο με το άθροισμα των στοιχείων του a^1 .

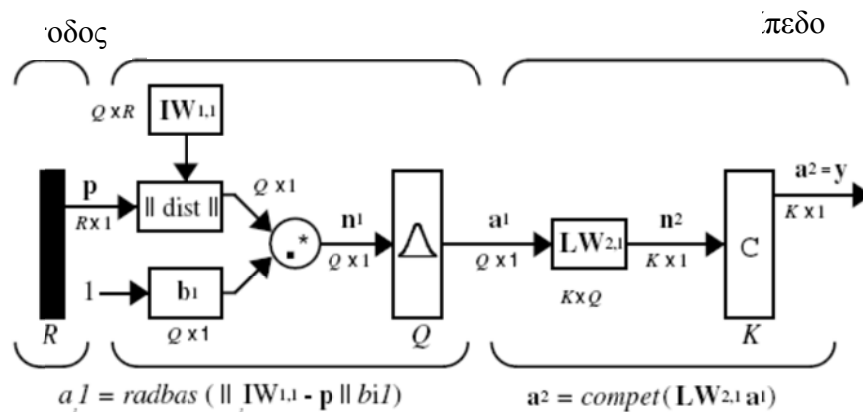


Σχήμα 3.7. Αρχιτεκτονική δικτύου GRNN [32].

Το κρυφό επίπεδο είναι ένα επίπεδο αποτελούμενο από RBF νευρώνες. Έχει τόσους νευρώνες όσο είναι το πλήθος των ζευγών των διανυσμάτων εισόδου/ εξόδου (Q). Επίσης, ο αριθμός νευρώνων στο επίπεδο εξόδου ισούται με το πλήθος των ζευγών των διανυσμάτων εισόδου / εξόδου (Q).

Για την υλοποίηση του GRNN, εκτός από τα διανύσματα εισόδου / εξόδου, απαιτείται μια τιμή spread. Η επιλογή της τιμής του spread είναι καθοριστική για την απόδοση του δικτύου. Μια μεγάλη τιμή spread οδηγεί σε μια μεγάλη περιοχή γύρω από το διάνυσμα εισόδου όπου οι νευρώνες του κρυφού επιπέδου θα ανταποκρίνονται με μεγάλες τιμές εξόδου. Όσο το spread μεγαλώνει, τόσο η κλίση της RBF γίνεται ομαλότερη και αρκετοί νευρώνες μπορεί να ανταποκρίνονται σε ένα διάνυσμα εξόδου. Αν η τιμή του spread είναι μικρή, τότε η RBF έχει απότομη κλίση, με αποτέλεσμα ο νευρώνας με το πλησιέστερο διάνυσμα βάρους στο διάνυσμα εισόδου να δώσει μια πολύ μεγαλύτερη έξοδο από τους άλλους νευρώνες.

Τα PNNs, που προσεγγίζουν την Bayesian στατιστική τεχνική, συνδυάζουν νέα διανύσματα εισόδου με υπάρχοντα αποθηκευμένα δεδομένα με σκοπό την ορθή ταξινόμηση των δεδομένων εισόδου [8]. Αυτός ο τρόπος λειτουργίας τους μοιάζει με την ανθρώπινη συμπεριφορά [38]. Τα PNNs βασίζονται στη συνάρτηση πυκνότητας πιθανότητας (Probabilistic Density Function - PDF) του Parzen [39] και αποτελούνται από ένα RBF κρυφό επίπεδο και από ένα ανταγωνιστικό επίπεδο (competitive layer) εξόδου [4, 6, 8, 32]. Όταν στο επίπεδο εισόδου ενός PNN παρουσιαστεί ένα διάνυσμα εισόδου, το κρυφό επίπεδο υπολογίζει την απόσταση μεταξύ του διανύσματος εισόδου και των διανυσμάτων εκπαίδευσης και παράγει ένα διάνυσμα που τα στοιχεία του δείχνουν πόσο κοντά είναι το διάνυσμα εισόδου σε κάθε διάνυσμα εκπαίδευσης. Το ανταγωνιστικό επίπεδο αθροίζει αυτές τις αποστάσεις και παράγει ένα διάνυσμα πιθανοτήτων. Τελικά, μια ανταγωνιστική συνάρτηση εξόδου (compete output function) βρίσκει τη μέγιστη από τις πιθανότητες και παράγει 1 για το αντίστοιχο στοιχείο και 0 για όλα τα άλλα. Η αρχιτεκτονική ενός PNN παρουσιάζεται στο Σχήμα 3.8. Τα PNNs χρησιμοποιούνται σε προβλήματα ταξινόμησης.



Σχήμα 3.8. Αρχιτεκτονική PNN [32].

Ο αριθμός των νευρώνων εισόδου ενός PNN ισούται με τον αριθμό των μεταβλητών του προβλήματος κι ο αριθμός των νευρώνων εξόδου προκύπτει από το πλήθος (K) των κλάσεων στις

οποίες μπορούν να ταξινομηθούν τα διανύσματα εισόδου. Ο αριθμός των νευρώνων στο κρυφό επίπεδο ισούται με το πλήθος των προτύπων εισόδου/εξόδου (Q) που χρησιμοποιήθηκαν κατά τη δημιουργία του PNN [8].

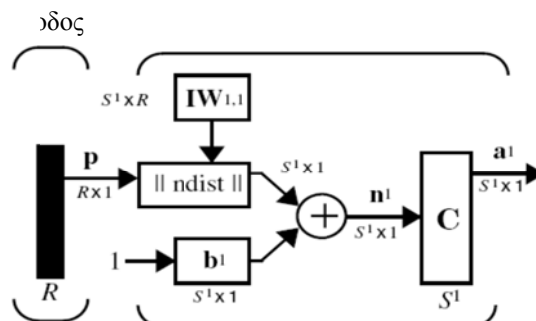
Τα PNNs δεν απαιτούν επαναληπτική διαδικασία εκπαίδευσης, με αποτέλεσμα να μπορούν να επεξεργαστούν μεγαλύτερο πλήθος δεδομένων σε μικρότερο χρόνο από τα MLPs. Η ιδιότητα αυτή των πιθανοκρατικών νευρωνικών δικτύων προκύπτει από τη συμπεριφορά της Bayesian στατιστικής τεχνικής [7].

Τα GRNNs και PNNs έχουν ένα σημαντικό μειονέκτημα. Είναι πιο αργά στη λειτουργία τους επειδή εκτελούν περισσότερους υπολογισμούς από άλλα είδη νευρωνικών δικτύων για να εκτελέσουν τις λειτουργίες προσέγγισης και ταξινόμησης [4, 32, 40].

3.5.6 Αυτο-Οργανώμενα Δίκτυα

Τα *αυτο-οργανώμενα νευρωνικά δίκτυα (Self-Organizing networks)* είναι ένα εξαιρετικού ενδιαφέροντος θέμα στο πεδίο των νευρωνικών δικτύων. Αυτά τα δίκτυα μπορούν να μαθαίνουν να ανιχνεύουν κανονικότητες και συσχετίσεις των εισόδων τους και να προσαρμόζουν την απόκρισή τους ανάλογα με τις εισόδους τους. Αυτο-οργανώμενα νευρωνικά δίκτυα είναι τα *ανταγωνιστικά δίκτυα (competitive networks)*, οι *αυτο-οργανώμενοι χάρτες χαρακτηριστικών (Self-Organizing Feature Maps - SOFMs)* και τα *δίκτυα γραμμικής κβαντοποίησης διανυσμάτων (Learning Vector Quantization – LVQ)*.

Ένα *ανταγωνιστικό δίκτυο* μαθαίνει να αναγνωρίζει ομάδες όμοιων διανυσμάτων εισόδου. Η αρχιτεκτονική ενός ανταγωνιστικού δικτύου παρουσιάζεται στο Σχήμα 3.9. Το $\|dist\|$ λαμβάνει το διάνυσμα εισόδου \mathbf{p} και τον πίνακα \mathbf{W} και παράγει ένα διάνυσμα \mathbf{S} στοιχείων. Τα στοιχεία του διανύσματος είναι η αρνητική τιμή των αποστάσεων μεταξύ του πίνακα \mathbf{W} και του διανύσματος εισόδου.



Σχήμα 3.9. Αρχιτεκτονική ανταγωνιστικού δικτύου [32].

Η είσοδος \mathbf{n} στη συνάρτηση μεταφοράς του ανταγωνιστικού επιπέδου υπολογίζεται από την εύρεση της αρνητικής απόστασης μεταξύ του πίνακα και του διανύσματος εισόδου, προσθέτοντας και τις αντίστοιχες πολώσεις \mathbf{b} . Αν όλες οι πολώσεις είναι μηδέν, τότε η μέγιστη τιμή που μπορεί να έχει η είσοδος είναι 0. Αυτό συμβαίνει όταν το διάνυσμα εισόδου \mathbf{p} ισούται με τις τιμές μιας από τις γραμμές του πίνακα \mathbf{W} .

Η ανταγωνιστική συνάρτηση εξόδου λαμβάνει το διάνυσμα εισόδου της και επιστρέφει σαν αποτέλεσμα 0 για όλους τους νευρώνες εκτός από το *νικητή (winner)*, που είναι ο νευρώνας που συνδέεται με το πιο θετικό στοιχείο του διανύσματος \mathbf{n} . Η έξοδος του νικητή ισούται με ένα. Αν όλες οι πολώσεις είναι μηδέν, ο νευρώνας που το διάνυσμα βάρους του είναι πλησιέστερο του διανύσματος εισόδου έχει τη *μικρότερη* αρνητική τιμή και συνεπώς κερδίζει τον ανταγωνισμό δίνοντας στην έξοδο του 1.

Οι *αυτο-οργανώμενοι χάρτες χαρακτηριστικών (Self-Organizing Feature Maps - SOFMs)* μαθαίνουν να ταξινομήσουν τα διανύσματα εισόδου σύμφωνα με τον τρόπο που είναι ομαδοποιημένα στο χώρο εισόδου (input space). Τα SOFMs διαφέρουν από τα ανταγωνιστικά επίπεδα (competitive layers) διότι οι γειτονικοί νευρώνες ενός SOFM μαθαίνουν να αναγνωρίζουν γειτονικά τμήματα του χώρου εισόδου. Επομένως τα SOFMs αποκτούν γνώση τόσο για την κατανομή (κάτι που κάνουν και τα ανταγωνιστικά επίπεδα) όσο και για την τοπολογία των διανυσμάτων εισόδου με τα οποία εκπαιδεύτηκαν, με αποτέλεσμα οι νευρώνες που είναι κοντά ο ένας με τον άλλο να ανταποκρίνονται σε όμοια διανύσματα εισόδου [6, 32, 34, 40].

Οι νευρώνες στο επίπεδο του SOFM διευθετούνται αρχικά σε φυσικές θέσεις σύμφωνα με μια συνάρτηση τοπολογίας (topology function). Αυτές οι συναρτήσεις μπορούν να ταξινομήσουν τους νευρώνες σε πλέγμα (grid), εξαγωνική (hexagonal) ή τυχαία (random) τοπολογία. Οι αποστάσεις μεταξύ των νευρώνων υπολογίζονται από τις θέσεις τους με μια συνάρτηση απόστασης (distance function). Οι συναρτήσεις απόστασης υπολογίζουν τις ακόλουθες αποστάσεις για δύο διανύσματα $\mathbf{P}_i = [p_{i1} \ p_{i2}]$ και $\mathbf{P}_j = [p_{j1} \ p_{j2}]$:

Ευκλείδεια απόσταση

$$d = \sqrt{(p_{i1} - p_{j1})^2 + (p_{i2} - p_{j2})^2}$$

Απόσταση Manhattan

$$d = |p_{i1} - p_{j1}| + |p_{i2} - p_{j2}|$$

Απόσταση σύνδεσης (link distance)

$$d = \begin{cases} 0, & \text{αν } i = j \\ 1, & \text{αν } \sqrt{(p_{i1} - p_{j1})^2 + (p_{i2} - p_{j2})^2} \leq 1 \\ 2, & \text{αν υπάρχει } k : D_{ik} = D_{kj} = 1 \\ 3, & \text{αν υπάρχουν } k_1, k_2 : D_{ik_1} = D_{k_1k_2} = D_{k_2j} = 1 \\ N, & \text{αν υπάρχουν } k_1 \dots k_N : D_{ik_1} = D_{k_1k_2} = \dots = D_{k_Nj} = 1 \\ S, & \text{όταν δεν πληρείται καμία από τις συνθήκες} \end{cases}$$

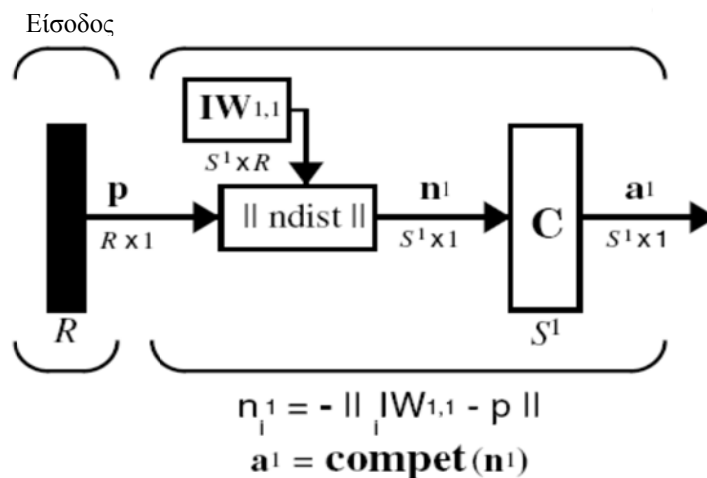
όπου $N \times S$ είναι οι διαστάσεις του πίνακα των θέσεων των νευρώνων, ενώ ο d έχει διαστάσεις $S \times S$.

Απόσταση box

$$d = \max(|p_{i1} - p_{j1}|, |p_{i2} - p_{j2}|)$$

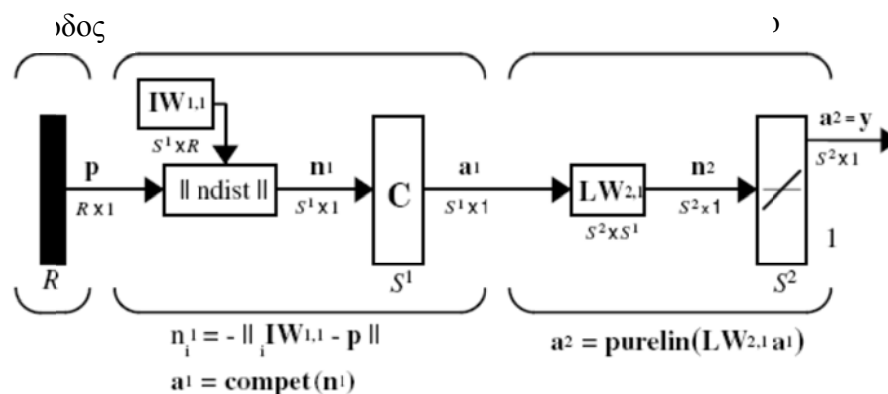
Ένας αυτο-οργανώμενος χάρτης χαρακτηριστικών προσδιορίζει το νευρώνα-νικητή χρησιμοποιώντας την ίδια διαδικασία με αυτή ενός ανταγωνιστικού επιπέδου. Όμως, αντί να τροποποιούνται μόνο τα βάρη του νευρώνα-νικητή, προσαρμόζονται τα βάρη όλων των νευρώνων που βρίσκονται σε μια προκαθορισμένη γειτονιά του νευρώνα-νικητή.

Η αρχιτεκτονική ενός SOFM παρουσιάζεται στο Σχήμα 3.103. Η αρχιτεκτονική μοιάζει με αυτήν ενός ανταγωνιστικού δικτύου, εκτός από το γεγονός ότι δεν υπάρχει πόλωση. Η ανταγωνιστική συνάρτηση μεταφοράς παράγει 1 στην έξοδο που αντιστοιχεί στο νευρώνα-νικητή.



Σχήμα 3.10. Αρχιτεκτονική SOFM [32].

Τα *Learning Vector Quantization (LVQ)* δίκτυα μαθαίνουν να ταξινομούν διανύσματα εισόδου σε κλάσεις-στόχους επιλεγμένες από το χρήστη. Η αρχιτεκτονική ενός LVQ δικτύου απεικονίζεται στο Σχήμα 3.11. Ένα LVQ δίκτυο έχει ένα ανταγωνιστικό κρυφό επίπεδο και ένα γραμμικό επίπεδο εξόδου. Το ανταγωνιστικό επίπεδο χρησιμοποιείται για την ταξινόμηση των διανυσμάτων εισόδου σε *υποκατηγορίες* (subclasses), ενώ το γραμμικό επίπεδο μετασχηματίζει τις κλάσεις του ανταγωνιστικού επιπέδου στις επιθυμητές *κλάσεις-στόχους* όπως καθορίζονται από το σχεδιαστή του δικτύου [5, 32, 41]. Οι κλάσεις που προκύπτουν από το ανταγωνιστικό επίπεδο ονομάζονται *υποκατηγορίες* (subclasses), ενώ αυτές του γραμμικού επιπέδου σαν *κλάσεις-στόχοι* (target classes). Τόσο στο ανταγωνιστικό όσο και στο γραμμικό επίπεδο ο αριθμός των νευρώνων προσδιορίζεται από τον αριθμό των υποκατηγοριών και των κλάσεων-στόχων, αντίστοιχα. Μάλιστα, ο αριθμός των νευρώνων του ανταγωνιστικού επιπέδου είναι πάντοτε μεγαλύτερος από το πλήθος των νευρώνων στο γραμμικό επίπεδο [32].



Σχήμα 3.11. Αρχιτεκτονική LVQ [32].

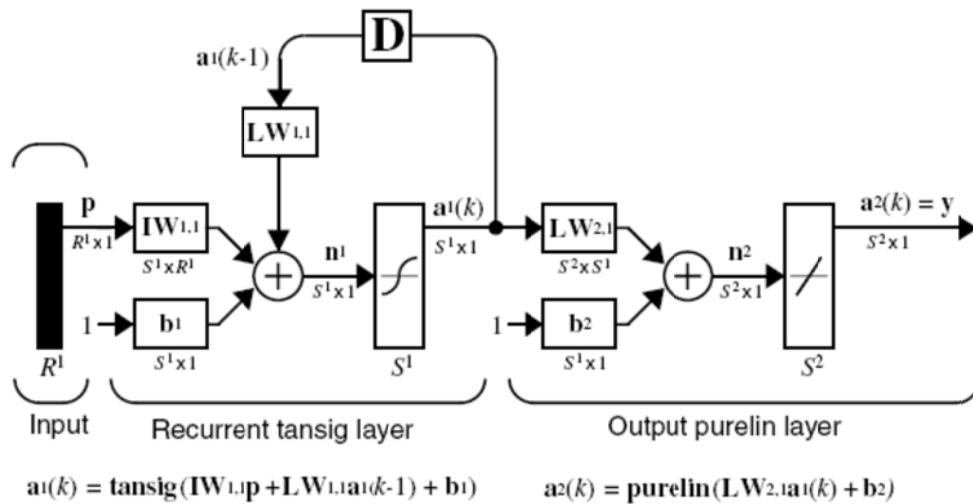
Τα LVQ δίκτυα μπορούν να ταξινομήσουν κάθε σύνολο διανυσμάτων εισόδου, χωρίς να απαιτείται τα διανύσματα να είναι γραμμικά διαχωρίσιμα. Η μόνη απαίτηση είναι το ανταγωνιστικό επίπεδο να έχει αρκετούς νευρώνες [4, 32].

3.5.7 Δίκτυο Elman

Το δίκτυο Elman είναι ένα τριών επιπέδων δίκτυο με ανάδραση από την έξοδο του κρυφού επιπέδου στην είσοδό του. Η ανάδραση επιτρέπει στο δίκτυο Elman να ανιχνεύει και να παράγει πρότυπα χρονικά μεταβαλλόμενα. Το δίκτυο Elman παρουσιάζεται στο Σχήμα 3.12.

Το δίκτυο Elman έχει νευρώνες με hyperbolic tangent sigmoid συναρτήσεις εξόδου στο κρυφό επίπεδο και γραμμικούς νευρώνες στο επίπεδο εξόδου. Αυτός ο συνδυασμός είναι ξεχωριστός καθώς τα δυο αυτά επίπεδα με τις συγκεκριμένες συναρτήσεις εξόδου μπορούν να

προσεγγίσουν κάθε συνάρτηση (με ένα πεπερασμένο αριθμό ασυνεχειών) με όση ακρίβεια απαιτείται από την εφαρμογή. Η μόνη απαίτηση είναι το κρυφό επίπεδο να έχει όσο περισσότερους νευρώνες είναι εφικτό.



Σχήμα 3.12. Αρχιτεκτονική δικτύου Elman [32].

Η διαφορά του δικτύου Elman από ένα τυπικό δίκτυο τριών επιπέδων έγκειται στο γεγονός ότι το κρυφό επίπεδο έχει ανάδραση. Η καθυστέρηση αυτή αποθηκεύει τιμές από προηγούμενα χρονικά βήματα, που μπορούν να χρησιμοποιηθούν στο τρέχον χρονικό βήμα. Επομένως, δυο δίκτυα Elman με τα ίδια βάρη και πολώσεις και με τις ίδιες εισόδους σε μια χρονική στιγμή, μπορεί να δώσουν διαφορετικές εξόδους εξαιτίας των διαφορετικών καταστάσεων ανάδρασης.

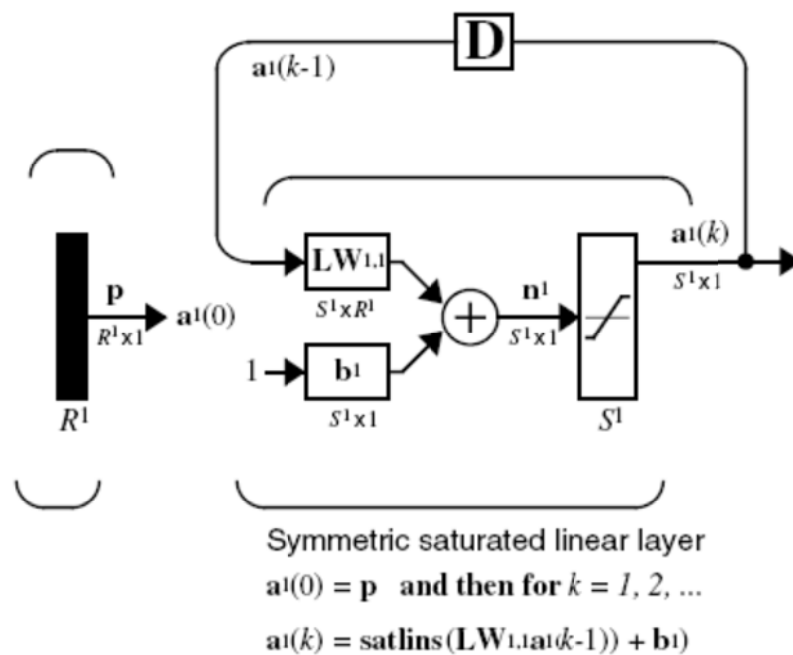
Επειδή το δίκτυο μπορεί να αποθηκεύει πληροφορίες για μελλοντική χρήση, είναι ικανό να μαθαίνει τόσο χρονικά όσο και χωρικά πρότυπα. Επομένως, το δίκτυο Elman είναι χρήσιμο στην επεξεργασία σήματος καθώς και σε εφαρμογές πρόβλεψης, όπου ο χρόνος διαδραματίζει πρωτεύοντα ρόλο [4].

3.5.8 Δίκτυο Hopfield

Ο σκοπός ενός δικτύου Hopfield είναι να αποθηκεύει ένα συγκεκριμένο σύνολο διανυσμάτων αναφοράς, ώστε όταν εισαχθεί μια νέα κατάσταση στο δίκτυο να σταθεροποιηθεί σε ένα διάνυσμα αναφοράς. Τα διανύσματα αναφοράς μπορούν να θεωρηθούν σαν μνήμες που το δίκτυο ανακαλεί όταν τροφοδοτείται με παρόμοια διανύσματα. Το δίκτυο Hopfield είναι αναδρομικό καθώς η έξοδος τροφοδοτείται στην είσοδο. Ένα δίκτυο Hopfield μπορεί να λειτουργήσει σαν δίκτυο αντιστάθμισης σφάλματος (error correction) ή ταξινόμησης διανυσμάτων

(vectors categorization) [4, 32, 40]. Η αρχιτεκτονική ενός δικτύου Hopfield παρουσιάζεται στο Σχήμα 3.13.

Το δίκτυο Hopfield χρησιμοποιεί τη συμμετρική κορεσμένη γραμμική συνάρτηση (symmetric saturating linear function). Το δίκτυο ελέγχεται με ένα ή περισσότερα διανύσματα εισόδου που παρουσιάζονται σαν αρχικές συνθήκες στο δίκτυο. Μετά την εισαγωγή των αρχικών συνθηκών, το δίκτυο παράγει μια έξοδο που τροφοδοτείται σαν είσοδος στο δίκτυο. Η διαδικασία αυτή επαναλαμβάνεται μέχρι να σταθεροποιηθούν οι έξοδοι. Είναι επιθυμητό το διάνυσμα εξόδου τελικά να συγκλίνει σε ένα από τα σχεδιασμένα διανύσματα αναφοράς που είναι πλησιέστερο στην είσοδο που το διεγείρει [32].



Σχήμα 3.13. Αρχιτεκτονική δικτύου Hopfield [32].

3.6 Εκπαίδευση Τεχνητού Νευρωνικού Δικτύου

Αντικείμενο ενός ΤΝΔ είναι να πραγματοποιεί μια συγκεκριμένη λειτουργία, όταν δεχθεί στις εισόδους του ένα σύνολο δεδομένων. Η διαδικασία που διασφαλίζει ότι το ΤΝΔ υπολογίζει τη σωστή έξοδο για όλες τις γνωστές καταστάσεις ονομάζεται **μάθηση (learning)** ή **εκπαίδευση (training)** του νευρωνικού δικτύου [4, 35] και βασίζεται σε **κανόνες εκμάθησης (learning rules)** [32]. Η εκπαίδευση ενός νευρωνικού δικτύου αναφέρεται στην προσαρμογή των τιμών των βαρών (μετακίνηση υπερεπιπέδων (hyperplanes)/ υπερσφαιρών (hyperspheres)) μεταξύ των νευρώνων [42, 43], ώστε ένα συγκεκριμένο διάνυσμα εισόδου (input vector) να δίνει μια επιθυμητή τιμή στην έξοδο (target output). Εντούτοις, ορισμένες φορές, η εκπαίδευση περιλαμβάνει την τροποποίηση

της τοπολογίας του δικτύου, προσθέτοντας ή αφαιρώντας συνδέσεις μεταξύ των νευρώνων του δικτύου. Η τροποποίηση των βαρών είναι πιο γενική από τη μεταβολή της τοπολογίας, καθώς ένα δίκτυο με πάρα πολλές συνδέσεις μπορεί να εκπαιδευτεί ώστε να θέτει το βάρος καθεμίας στο μηδέν, που ισοδυναμεί με την εξάλειψη της συγκεκριμένης σύνδεσης.

Υπάρχουν δυο τύποι μάθησης, η *επιβλεπόμενη μάθηση* (*supervised learning*) και η *μη-επιβλεπόμενη μάθηση* (*unsupervised learning*). Η επιβλεπόμενη μάθηση περιλαμβάνει ένα μηχανισμό τροφοδότησης του δικτύου τόσο με τα διανύσματα εισόδου όσο και με τις επιθυμητές εξόδους κατά τη φάση εκπαίδευσης (training phase). Η μη-επιβλεπόμενη μάθηση χρησιμοποιεί μόνο τα διανύσματα εισόδου για να ανιχνεύσει τις μεταξύ τους κανονικότητες, χωρίς εξωτερική βοήθεια [4, 40]. Η πλειοψηφία των νευρωνικών δικτύων βασίζεται στην επιβλεπόμενη μάθηση, ενώ η μη-επιβλεπόμενη εκπαίδευση χρησιμοποιείται για την εκτέλεση ενός αρχικού διαχωρισμού των εισόδων.

3.6.1 Επιβλεπόμενη Μάθηση

Η επιβλεπόμενη μάθηση ενσωματώνει έναν “εξωτερικό δάσκαλο” (external teacher), ώστε σε κάθε νευρώνα εξόδου να είναι γνωστή η επιθυμητή απόκριση για τα σήματα εισόδου. Για το σκοπό αυτό το νευρωνικό δίκτυο δέχεται τόσο τα διανύσματα εισόδου, όσο και τις αντίστοιχες επιθυμητές εξόδους. Το δίκτυο κατά τη φάση της εκπαίδευσής του, επεξεργάζεται τα διανύσματα εισόδου και συγκρίνει τις πραγματικές (actual) αποκρίσεις με τις επιθυμητές εξόδους. Αν οι τιμές αυτές διαφέρουν, τα βάρη προσαρμόζονται ώστε στην επόμενη επανάληψη οι επιθυμητές και οι πραγματικές εξοδοί να έχουν μικρότερη απόκλιση. Η επιβλεπόμενη μάθηση προσπαθεί να ελαχιστοποιήσει το σφάλμα (error) που παράγεται από τους νευρώνες. Η διαδικασία επαναλαμβάνεται μέχρι να επιτευχθεί μια αποδεκτή ακρίβεια (accuracy) του δικτύου.

Με την επιβλεπόμενη μάθηση, το νευρωνικό δίκτυο πρέπει πρώτα να εκπαιδευτεί πριν χρησιμοποιηθεί σε κάποια εφαρμογή. Το σύνολο των δεδομένων που χρησιμοποιείται κατά τη διάρκεια της εκπαίδευσης ονομάζεται “*σύνολο εκπαίδευσης*” (*training set*). Κατά τη φάση εκπαίδευσης ενός νευρωνικού δικτύου το ίδιο σύνολο δεδομένων επεξεργάζεται από το ΤΝΔ πολλές φορές μέχρι τα συναπτικά βάρη να βελτιστοποιηθούν [40].

Η εκπαίδευση θεωρείται ολοκληρωμένη όταν το νευρωνικό δίκτυο φτάσει το επίπεδο απόδοσης που έχει ορίσει ο χρήστης. Το επίπεδο αυτό φανερώνει ότι το νευρωνικό δίκτυο πέτυχε την επιθυμητή στατιστική ακρίβεια και παράγει τις κατάλληλες εξόδους για μια ακολουθία δεδομένων εισόδων. Όταν πλέον δεν χρειάζεται περαιτέρω εκπαίδευση, τα βάρη παραμένουν αμετάβλητα για την εφαρμογή.

Όταν τα δεδομένα εισόδου δεν περιέχουν πληροφορίες που να συσχετίζονται με την επιθυμητή έξοδο, τότε το ΤΝΔ δεν μπορεί να μάθει [4, 40]. Επίσης, το νευρωνικό δίκτυο δεν συγκλίνει εάν τα δεδομένα δεν είναι αρκετά, ώστε να υπάρξει πλήρης εκπαίδευσή του.

Στην επιβλεπόμενη μάθηση μπορούμε να διακρίνουμε δυο τύπους εκπαίδευσης, την *εκπαίδευση κατά δέσμες (batch training)* και τη *σταδιακή εκπαίδευση (incremental training)* [32, 40]. Στον πρώτο τύπο μάθησης, τα συναπτικά βάρη τροποποιούνται βάση μιας δέσμης διανυσμάτων εισόδου. Η εκπαίδευση κατά δέσμες ονομάζεται και off-line μάθηση. Στη σταδιακή εκπαίδευση τα βάρη προσαρμόζονται κάθε φορά που ένα νέο διάνυσμα εισόδου εισαχθεί στο νευρωνικό δίκτυο. Η σταδιακή εκπαίδευση είναι γνωστή και σαν on-line ή προσαρμόσιμη εκπαίδευση [4, 32].

3.6.2 Μη-Επιβλεπόμενη Μάθηση

Στη μη-επιβλεπόμενη μάθηση δεν υπάρχει “εξωτερικός δάσκαλος”. Το νευρωνικό δίκτυο πρέπει μόνο του να ανιχνεύσει τις κανονικότητες (regularities) των δεδομένων εισόδου του. Η παρακολούθηση της απόδοσης του νευρωνικού δικτύου γίνεται εσωτερικά. Το δίκτυο τροφοδοτείται μόνο με τα δεδομένα εισόδου, χωρίς τις επιθυμητές εξόδους. Η μη-επιβλεπόμενη μάθηση, συχνά, αναφέρεται σαν αυτο-οργάνωση (self-organization) ή προσαρμογή (adaptation).

Επειδή το νευρωνικό δίκτυο δεν έχει κάποιο μηχανισμό για να ελέγξει την ορθότητα των αποτελεσμάτων του, χρειάζεται να τροφοδοτηθεί με πληροφορίες σχετικά με τον τρόπο οργάνωσής του. Αυτές οι πληροφορίες βασίζονται στην τοπολογία του δικτύου και τους κανόνες εκμάθησης.

3.7 Κανόνες Εκμάθησης Τεχνητού Νευρωνικού Δικτύου

Για την εκπαίδευση ενός ΤΝΔ ώστε να επιτελεί τη λειτουργία για την οποία κατασκευάστηκε, απαιτείται πέραν του τύπου μάθησης, η χρησιμοποίηση κάποιου *κανόνα εκμάθησης (learning rule)*. Ο κανόνας εκμάθησης αναφέρεται επίσης σαν *αλγόριθμος εκπαίδευσης (training algorithm)* και είναι η διαδικασία προσαρμογής των βαρών ενός νευρωνικού δικτύου. Υπάρχουν πολλοί αλγόριθμοι εκπαίδευσης στη βιβλιογραφία [32]. Μια συνοπτική περιγραφή δίνεται παρακάτω.

3.7.1 Κανόνας του Hebb (Hebb's Rule)

Ο πρώτος και αναμφίβολα ο γνωστότερος κανόνας εκμάθησης παρουσιάστηκε από τον Donald Hebb (*Κανόνας του Hebb - Hebb's Rule*). Η περιγραφή υπάρχει στο βιβλίο του “The Organization of Behavior” που εκδόθηκε το 1949. Σύμφωνα με τον κανόνα του Hebb, αν ένας νευρώνας λάβει ένα σήμα εισόδου από άλλο νευρώνα, κι αν και οι δύο νευρώνες είναι ενεργοποιημένοι (μαθηματικά έχουν το ίδιο πρόσημο) το βάρος μεταξύ των νευρώνων ισχυροποιείται [43].

3.7.2 Αλγόριθμος Ελάχιστου Μέσου Τετραγωνικού Σφάλματος

Ο αλγόριθμος ελάχιστου μέσου τετραγωνικού σφάλματος (*Least Mean Square error Algorithm – LMS*) ονομάζεται, επίσης, κανόνας Δέλτα (Delta rule) ή κανόνας των Widrow-Hoff [4] και είναι παραλλαγή του κανόνα του Hebb [40]. Ο αλγόριθμος βασίζεται στην ιδέα της συνεχούς τροποποίησης των βαρών του νευρωνικού δικτύου για τη μείωση της διαφοράς (delta) μεταξύ των τιμών των επιθυμητών και πραγματικών εξόδων. Ο αλγόριθμος μεταβάλλει τα βάρη ώστε να ελαχιστοποιηθεί το μέσο τετραγωνικό σφάλμα του δικτύου. Ο LMS χρησιμοποιείται για την εκπαίδευση νευρωνικών δικτύων με γραμμικές συναρτήσεις ενεργοποίησης.

Οι Widrow-Hoff εκτίμησαν το μέσο τετραγωνικό σφάλμα χρησιμοποιώντας το τετραγωνικό σφάλμα κάθε επανάληψης. Ο υπολογισμός της μερικής παραγώγου του τετραγωνικού σφάλματος ως προς τα βάρη και τις πολώσεις στην k επανάληψη δίνει

$$\frac{\partial e_k^2}{\partial w_{1,j}} = 2e_k \frac{\partial e_k}{\partial w_{1,j}}, \text{ για } j = 1, 2, \dots, R$$

και

$$\frac{\partial e_k^2}{\partial b} = 2e_k \frac{\partial e_k}{\partial b}$$

Η μερική παράγωγος του σφάλματος υπολογίζεται ως εξής

$$\frac{\partial e_k}{\partial w_{1,j}} = \frac{\partial [t_k - a_k]}{\partial w_{1,j}} = \frac{\partial}{\partial w_{1,j}} [t_k - (Wp_k + b)] \Rightarrow$$

$$\frac{\partial e_k}{\partial w_{1,j}} = \frac{\partial}{\partial w_{1,j}} \left[t_k - \sum_{i=1}^R (w_{1,i} p_k(i) + b) \right]$$

Το $p_k(i)$ είναι το i στοιχείο του διανύσματος εισόδου στην k επανάληψη.

Συνεπώς,

$$\frac{\partial e_k}{\partial w_{1,j}} = -p_k(j)$$

Ομοίως

$$\frac{\partial e_k}{\partial b} = -1$$

Επομένως, η προσαρμογή των βαρών και των πόλωσης θα δίνεται από τις ποσότητες $2ae_k p_k$ και $2ae_k$. Αυτές οι δυο εξισώσεις περιγράφουν τον αλγόριθμο εκμάθησης των Widrow-Hoff. Η γενίκευση των εξισώσεων αυτών στην περίπτωση πολλαπλών επιπέδων νευρώνων είναι οι ακόλουθες μαθηματικές σχέσεις.

$$W_{k+1} = W_k + 2ae_k p_k$$

$$b_{k+1} = b_k + 2ae_k$$

Εδώ το σφάλμα e και η πόλωση b είναι διανύσματα και το a είναι ο ρυθμός εκμάθησης. Ο ρυθμός εκμάθησης επιδρά στην ταχύτητα με την οποία μαθαίνει το νευρωνικό δίκτυο και καθορίζει τη σύγκλιση του αλγορίθμου εκπαίδευσης. Ένας μικρός ρυθμός εκμάθησης απαιτεί περισσότερο χρόνο για την πραγματοποίηση της εκπαίδευσης και τη δημιουργία ενός ικανοποιητικά εκπαιδευμένου δικτύου. Ένας υψηλότερος ρυθμός εκμάθησης θα επιτάχυνε τη διαδικασία εκπαίδευσης, αλλά το δίκτυο δεν θα πετύχαινε τις καλύτερες δυνατές διακρίσεις των καταστάσεων των δεδομένων.

Ο τρόπος που λειτουργεί ο αλγόριθμος LMS είναι ότι το σφάλμα στο επίπεδο εξόδου μεταφέρεται μέσω της παραγώγου της συνάρτησης ενεργοποίησης και χρησιμοποιείται στο προηγούμενο επίπεδο για να προσαρμόσει τα βάρη στο επίπεδο εισόδου. Με άλλα λόγια, το σφάλμα μεταδίδεται προς τα πίσω στα προηγούμενα επίπεδα, με ένα επίπεδο κάθε φορά. Η διαδικασία της οπισθοδιάδοσης (backpropagation) του σφάλματος του δικτύου συνεχίζεται μέχρι να φτάσει στο πρώτο επίπεδο.

3.7.3 Αλγόριθμος Πτωτικής Κλίσης

Ο *αλγόριθμος πτωτικής κλίσης (gradient descent algorithm)* χρησιμοποιείται ευρέως για την εκπαίδευση των νευρωνικών δικτύων. Στόχος του είναι η προσαρμογή των βαρών ενός νευρωνικού δικτύου ώστε το σφάλμα του δικτύου να μειώνεται. Για την εύρεση της φοράς μείωσης του σφάλματος χρησιμοποιείται η πρώτη παράγωγος ή κλίση του. Η τροποποίηση των βαρών του δικτύου γίνεται βάση της μαθηματικής σχέσης

$$W_{k+1} = W_k - a_k g_k$$

όπου W_k είναι το διάνυσμα των βαρών και των πολώσεων, g_k η τρέχουσα κλίση και a_k ο ρυθμός εκμάθησης (learning rate).

Ο αλγόριθμος πτωτικής κλίσης ονομάζεται και steepest descent. Στους αλγορίθμους πτωτικής κλίσης υπάγεται ο αλγόριθμος οπισθοδιάδοσης, που αναλύεται στη συνέχεια [32].

3.7.4 Αλγόριθμος Οπισθοδιάδοσης

Ο *αλγόριθμος οπισθοδιάδοσης (backpropagation algorithm)* είναι ένας από τους πιο δημοφιλείς αλγορίθμους για την εκπαίδευση ενός νευρωνικού δικτύου λόγω της απλότητας και της καταλληλότητάς του. Ο αλγόριθμος αυτός αποτελεί γενίκευση του κανόνα εκμάθησης των Widrow-Hoff και ανήκει στην οικογένεια των αλγορίθμων πτωτικής κλίσης (gradient descent). Ο αλγόριθμος αποτελείται από δυο φάσεις: *τη φάση εκπαίδευσης (training phase)* και τη *φάση ανάκλησης (recall phase)*. Στη φάση εκπαίδευσης, πρώτα, τα βάρη του δικτύου αρχικοποιούνται με τυχαίο τρόπο. Έπειτα, το νευρωνικό δίκτυο υπολογίζει την πραγματική έξοδο, ενός διανύσματος εισόδου, και τη συγκρίνει με την επιθυμητή τιμή. Στη συνέχεια, υπολογίζεται το σφάλμα και χρησιμοποιείται για την προσαρμογή των βαρών του επιπέδου εξόδου. Με παρόμοιο τρόπο, το σφάλμα διαδίδεται προς τα πίσω και χρησιμοποιείται για να προσαρμόσει τα βάρη του προηγούμενου επιπέδου.

Στη φάση ανάκλησης, μόνον οι προς τα εμπρός (feed-forward) υπολογισμοί, χρησιμοποιώντας τα προσαρμοσμένα βάρη από τη φάση εκπαίδευσης και τα διανύσματα εισόδου, λαμβάνουν χώρα. Η προς τα εμπρός διαδικασία χρησιμοποιείται τόσο στη φάση ανάκλησης όσο και στη φάση εκπαίδευσης. Από την άλλη πλευρά, η οπισθοδιάδοση του σφάλματος αξιοποιείται μόνο στην φάση εκπαίδευσης.

Στη φάση εκπαίδευσης, πρώτα αρχικοποιούνται τα βάρη του δικτύου. Έπειτα, υπολογίζεται η έξοδος κάθε επιπέδου, ξεκινώντας από το επίπεδο εισόδου και καταλήγοντας στο επίπεδο εξόδου. Εν συνεχεία, υπολογίζεται το σφάλμα για το επίπεδο εξόδου μέσω της σύγκρισης της πραγματικής εξόδου και της επιθυμητής τιμής για την προσαρμογή των βαρών στο επίπεδο εξόδου και στο κρυφό επίπεδο [44]. Η προσαρμογή των βαρών του δικτύου γίνεται σύμφωνα με την ακόλουθη σχέση

$$W_{k+1} = W_k - a_k g_k$$

όπου W_k είναι το διάνυσμα των βαρών και των πολώσεων, g_k η τρέχουσα κλίση και a_k ο ρυθμός εκμάθησης (learning rate).

3.7.5 Αλγόριθμος Πτωτικής Κλίσης με Ροπή

Ο αλγόριθμος πτωτικής κλίσης με ροπή (*gradient descent with momentum algorithm*) δουλεύει με παρόμοιο τρόπο με τον αλγόριθμο *gradient descent*, αλλά συγκλίνει γρηγορότερα. Η ροπή (*momentum*) παίρνει τιμές στο διάστημα $[0, 1]$. Όταν η τιμή της ροπής είναι κοντά στο μηδέν, η αλλαγή των βαρών βασίζεται στην πτωτική κλίση. Αντίθετα, όταν η ροπή πάρει τιμή κοντά στο 1, η προσαρμογή των βαρών επηρεάζεται σε μεγάλο βαθμό από την τελευταία τους αλλαγή και η κλίση αγνοείται. Η ροπή επιτρέπει στο νευρωνικό δίκτυο να αγνοεί ανεπαίσθητες μεταβολές του σφάλματος. Χωρίς τη ροπή, ένα νευρωνικό δίκτυο μπορεί να εγκλωβιστεί σε ένα ρηχό τοπικό ελάχιστο, ενώ με τη ροπή, το δίκτυο μπορεί να αποφύγει ένα τέτοιο ελάχιστο [32].

3.7.6 Πτωτική Κλίση με Μεταβλητό Ρυθμό Εκμάθησης (*Gradient Descent with Variable Learning Rate*)

Στον κανονικό αλγόριθμο πτωτικής κλίσης (*gradient descent*) ο ρυθμός εκμάθησης παραμένει σταθερός σε όλη τη διάρκεια της εκπαίδευσης. Η απόδοση ενός αλγορίθμου επηρεάζεται από την επιλογή του κατάλληλου ρυθμού εκπαίδευσης. Αν η τιμή του ρυθμού εκμάθησης είναι μεγάλη, ο αλγόριθμος πιθανόν να ταλαντώνεται και να γίνει ασταθής. Από την άλλη πλευρά, αν η τιμή του είναι πολύ μικρή, θα απαιτηθεί αρκετός χρόνος μέχρι ο αλγόριθμος να συγκλίνει. Η βέλτιστη τιμή του ρυθμού εκμάθησης μεταβάλλεται κατά τη διάρκεια της φάσης εκπαίδευσης, καθώς αλλάζει η απόδοση του δικτύου.

Η απόδοση του *steepest descent* αλγορίθμου μπορεί να βελτιωθεί, αν ο ρυθμός εκμάθησης μεταβάλλεται κατά τη διάρκεια της εκπαίδευσης, κάτι που εξασφαλίζεται από τον αλγόριθμο *πτωτικής κλίσης με μεταβλητό ρυθμό εκμάθησης (gradient descent with variable learning rate)*. Ένας προσαρμόσιμος ρυθμός εκμάθησης προσπαθεί να κρατήσει το μέγεθος του βήματος εκμάθησης όσο το δυνατόν μεγαλύτερο, ενώ παράλληλα, διατηρείται σταθερή η εκπαίδευση. Πρώτα, ο ρυθμός εκμάθησης αρχικοποιείται με μια συγκεκριμένη τιμή (*lr_init*). Αν το σφάλμα που έχει υπολογιστεί από το νευρωνικό δίκτυο, μειώνεται σταθερά, ο ρυθμός εκμάθησης αυξάνεται σύμφωνα με ένα σταθερό συντελεστή (*lr_inc*). Αυτό βοηθάει τον αλγόριθμο ώστε να συγκλίνει γρηγορότερα. Αν το σφάλμα αυξάνεται περισσότερο από ένα προκαθορισμένο ρυθμό (*max_perf_inc*), ο ρυθμός μετάδοσης μειώνεται με ένα σταθερό συντελεστή (*lr_decr*) [32].

3.7.7 Ευπροσάρμοστη Οπισθοδιάδοση

Τα MLPs στα κρυφά επίπεδα χρησιμοποιούν τις sigmoid και tansigmoid συναρτήσεις ενεργοποίησης. Οι συναρτήσεις αυτές αποκαλούνται συναρτήσεις “συμπίεσης” (*squashing functions*), διότι συμπιέζουν ένα απεριόριστο εύρος διανυσμάτων εισόδου σε ένα πεπερασμένο εύρος εξόδων. Οι *squashing* συναρτήσεις χαρακτηρίζονται από το γεγονός ότι η κλίση τους (ή η παράγωγός τους) τείνει στο μηδέν όταν οι είσοδοι παίρνουν μεγάλες τιμές. Αυτό αποτελεί

πρόβλημα όταν χρησιμοποιείται κάποιος αλγόριθμος *steepest descent* για την εκπαίδευση ενός MLP σε συνδυασμό με τις *squashing* συναρτήσεις, καθώς η κλίση έχει πολύ μικρό μέγεθος και ως εκ τούτου συμβαίνουν μικρές αλλαγές στα βάρη, ακόμη κι αν αυτά απέχουν πολύ από τις βέλτιστες τιμές τους, με αποτέλεσμα η σύγκλιση του αλγορίθμου να είναι πολύ αργή.

Ο σκοπός του *αλγορίθμου της ευπροσάρμοστης οπισθοδιάδοσης (resilient backpropagation)* είναι η εξάλειψη των προαναφερθέντων συνεπειών του πλάτους των παραγώγων των συναρτήσεων. Ο συγκεκριμένος αλγόριθμος προσαρμόζει τα βάρη του δικτύου βασιζόμενος μόνο στο πρόσημο της παραγώγου (*sign of the derivative*) χωρίς να λαμβάνει υπόψη την τιμή της. Η ένδειξη χρησιμοποιείται για να καθορίσει την κατεύθυνση προς την οποία θα γίνει η προσαρμογή. Αν για δυο συνεχόμενες επιτυχείς επαναλήψεις η παράγωγος της συνάρτησης σφάλματος που αφορά τα βάρη έχει το ίδιο πρόσημο, η τιμή των βαρών αυξάνεται κατά ένα συντελεστή *delt_inc*. Ενώ, στην περίπτωση που για δυο συνεχόμενες επιτυχείς επαναλήψεις το πρόσημο της παραγώγου δεν είναι το ίδιο, οι τιμές των βαρών μειώνονται κατά ένα συντελεστή *delt_dec*. Αν η παράγωγος ισούται με μηδέν, τότε οι τιμές των βαρών παραμένουν αμετάβλητες [32].

Ο αλγόριθμος ευπροσάρμοστης οπισθοδιάδοσης είναι γενικά πολύ γρηγορότερος από τον κλασικό αλγόριθμο πτωτικής κλίσης. Ένα σημαντικό χαρακτηριστικό του είναι ότι απαιτεί μέτρια αύξηση χωρητικότητας της μνήμης. Επιπρόσθετα, η αποθήκευση των προσαρμοσμένων τιμών για κάθε βάρος ισοδυναμεί με την αποθήκευση της παραγώγου [32].

3.7.8 Αλγόριθμος Συζευγμένης Κλίσης

Ο κλασικός αλγόριθμος οπισθοδιάδοσης προσαρμόζει τα βάρη προς την κατεύθυνση που εμφανίζει τη μεγαλύτερη μείωση στην τιμή (*steepest descent*) της συνάρτησης σφάλματος (*negative of the gradient*). Αυτή είναι η κατεύθυνση προς την οποία η συνάρτηση μειώνεται πιο γρήγορα. Αποδεικνύεται ότι παρόλο που η συνάρτηση μειώνεται πάρα πολύ γρήγορα κατά μήκος της αρνητικής κατεύθυνσης της παραγώγου, δεν είναι αναγκαία η επίτευξη της ταχύτερης σύγκλισης του αλγορίθμου. Στον *αλγόριθμο συζευγμένης κλίσης (Conjugate Gradient Algorithm)*, εκτελείται μια αναζήτηση προς συζευγμένες κατευθύνσεις, που γενικά συγκλίνουν γρηγορότερα σε σύγκριση με τις *steepest descent* κατευθύνσεις. Υλοποιήσεις του αλγορίθμου συζευγμένης κλίσης αποτελούν οι [32]:

1. Fletcher-Reeves Update,
2. Polak-Ribière Update,
3. Powell-Beale Restarts και
4. Scaled Conjugate Gradient.

3.7.9 Αλγόριθμος *Quasi-Newton*

Η μέθοδος του Newton είναι μια παραλλαγή του αλγορίθμου συζευγμένης κλίσης για γρήγορη βελτιστοποίηση. Το βασικό βήμα της μεθόδου Newton είναι

$$W_{k+1} = W_k - A_k^{-1} g_k$$

όπου W_k είναι το διάνυσμα των βαρών και των πολώσεων, g_k η τρέχουσα κλίση και A_k^{-1} ο πίνακας Hessian (δεύτερης τάξης παράγωγοι) της συνάρτησης σφάλματος.

Η μέθοδος Newton συχνά συγκλίνει γρηγορότερα συγκριτικά με τον αλγόριθμο συζευγμένης κλίσης. Εντούτοις, είναι πολύπλοκος και απαιτεί πολλούς υπολογιστικούς πόρους για τον υπολογισμό του πίνακα Hessian για τα εμπρόσθια διάδοσης νευρωνικά δίκτυα.

Υπάρχει μια κατηγορία αλγορίθμων που βασίζονται στη μέθοδο Newton, αλλά δεν απαιτούν τον υπολογισμό των δεύτερης τάξης παραγώγων. Αυτοί ονομάζονται *quasi-Newton* (ή *secant*) μέθοδοι. Αυτοί ενημερώνουν (update) μια προσέγγιση του πίνακα Hessian σε κάθε επανάληψη του αλγορίθμου. Η ενημέρωση (update) υπολογίζεται σαν συνάρτηση της κλίσης. Η quasi-Newton μέθοδος, που έχει την μεγαλύτερη επιτυχία σε δημοσιευμένες μελέτες, προέρχεται από τους Broyden, Fletcher, Goldfarb και Shanno (BFGS).

Ο BFGS αλγόριθμος απαιτεί περισσότερους υπολογισμούς σε κάθε επανάληψη και περισσότερη μνήμη από τις μεθόδους συζευγμένης κλίσης, όμως γενικά συγκλίνει σε λιγότερες επαναλήψεις. Ο προσεγγιστικός πίνακας Hessian που πρέπει να αποθηκευθεί, έχει $n \times n$ διαστάσεις, όπου το n ισούται με το πλήθος των βαρών του δικτύου. Για πολύ μεγάλα δίκτυα συνιστάται η χρήση του αλγορίθμου οπισθοδιάδοσης ή ενός αλγορίθμου συζευγμένης κλίσης. Για μικρότερου μεγέθους δίκτυα ο BFGS αλγόριθμος είναι αποτελεσματική μέθοδο εκπαίδευσης [32].

3.7.10 Αλγόριθμος *One Step Secant*

Ο BFGS αλγόριθμος απαιτεί περισσότερους υπολογισμούς και μνήμη σε κάθε επανάληψη συγκριτικά με τον αλγόριθμο συζευγμένης κλίσης, επομένως, απαιτείται η δημιουργία μιας *secant* προσέγγισης με λιγότερες απαιτήσεις σε υπολογισμούς και μνήμη. Η μέθοδος *one step secant* (*OSS*) είναι μια προσπάθεια γεφύρωσης μεταξύ του αλγορίθμου συζευγμένης κλίσης και του quasi-Newton (*secant*) αλγορίθμου. Η μέθοδος αυτή δεν αποθηκεύει ολόκληρο τον πίνακα Hessian, αλλά θεωρεί ότι ο προηγούμενος Hessian πίνακας ήταν ο πίνακας ομοιότητας (*identity matrix*). Αυτό έχει το πλεονέκτημα ότι η νέα κατεύθυνση αναζήτησης μπορεί να υπολογιστεί χωρίς να απαιτείται ο υπολογισμός του αντίστροφου πίνακα.

Ο OSS αλγόριθμος απαιτεί ελαφρώς περισσότερη μνήμη και υπολογισμούς σε κάθε επανάληψη από τον αλγόριθμο συζευγμένης κλίσης. Αυτός μπορεί να θεωρηθεί σαν ένας συμβιβασμός μεταξύ του quasi-Newton αλγορίθμου και του αλγορίθμου συζευγμένης κλίσης [32].

3.7.11 Αλγόριθμος *Levenberg-Marquardt*

Ο αλγόριθμος Levenberg-Marquardt, όπως και ο quasi-Newton, σχεδιάστηκε για να προσεγγίζουν δεύτερης τάξης παραγώγους χωρίς τον υπολογισμό του πίνακα Hessian. Όταν η συνάρτηση σφάλματος είναι της μορφής αθροίσματος τετραγώνων, τότε ο πίνακας Hessian προσεγγίζεται από τη σχέση

$$H = J^T J$$

και η κλίση δίνεται από τον τύπο

$$g = J^T e$$

όπου J είναι ο πίνακας Jacob (Jacobian matrix) που περιέχει παραγώγους πρώτης τάξης για το σφάλμα του δικτύου που αντιστοιχεί στα βάρη του και e είναι το διάνυσμα σφάλματος του δικτύου.

Ο Jacobian πίνακας μπορεί να υπολογιστεί μέσω μιας τυπικής τεχνικής οπισθοδιάδοσης που είναι λιγότερο πολύπλοκη από τον υπολογισμό του πίνακα Hessian. Η προσαρμογή των βαρών με τον αλγόριθμο των Levenberg-Marquardt δίνεται από τη σχέση:

$$W_{k+1} = W_k - \left[J^T J + \mu I \right]^{-1} J^T e$$

Όταν η βαθμωτή παράμετρος μ είναι μηδέν, τότε πρόκειται για τη μέθοδο Newton, χρησιμοποιώντας την προσέγγιση του πίνακα Hessian. Όταν το μ έχει μεγάλη τιμή, ο αλγόριθμος Levenberg-Marquardt μετατρέπεται σε πτωτικής κλίσης με μικρό βήμα. Το μ μειώνεται μετά από κάθε επιτυχή επανάληψη (μείωση του σφάλματος) και αυξάνει μόνο όταν ένα βήμα αυξάνει το σφάλμα. Επειδή η μέθοδος Newton είναι ταχύτερη και πιο ακριβής κοντά σε ένα ελάχιστο σφάλμα, σκοπός είναι ο αλγόριθμος Levenberg-Marquardt να μεταπέσει όσο το δυνατόν γρηγορότερα στον αλγόριθμο Newton.

Ο αλγόριθμος Levenberg-Marquardt φαίνεται ότι είναι η γρηγορότερη μέθοδος για την εκπαίδευση μέτριου μεγέθους εμπρόσθιας διάδοσης νευρωνικά δίκτυα (μέχρι αρκετές εκατοντάδες βάρη) [32].

3.7.12 Κανόνας Εκμάθησης Kohonen

Αυτός ο κανόνας εκπαίδευσης, που αναπτύχθηκε από τον Τευνο Kohonen (**Kohonen Learning Rule**), χαρακτηρίζεται από δυο βασικά στοιχεία που είναι ο κανόνας προσαρμογής των βαρών και η έννοια της “γειτονιάς” των νευρώνων στο επίπεδο εξόδου. Στη διαδικασία αυτή οι νευρώνες εξόδου συναγωνίζονται για την ευκαιρία να τροποποιήσουν τα βάρη τους. Ο νευρώνας με τη μεγαλύτερη τιμή εξόδου δηλώνεται σαν **νικητής** και έχει τη δυνατότητα να εμποδίσει τους ανταγωνιστές του καθώς επίσης και να διεγείρει τους γειτονικούς νευρώνες του. Μόνο ο νευρώνας-νικητής επιτρέπεται να δώσει μια έξοδο, ενώ ο νευρώνας-νικητής και οι γειτονικοί του επιτρέπεται να προσαρμόσουν τα βάρη των συνδέσεών τους. Η τροποποίηση των βαρών που ανήκουν σε μια γειτονιά δίνεται από την ακόλουθη αναδρομική σχέση:

$$W_{k+1} = W_k + q_k (P - W_k)$$

όπου q_k είναι ο ρυθμός εκπαίδευσης στην k επανάληψη που προσδιορίζει το μέγεθος της μεταβολής που θα γίνεται στα βάρη και συνεχώς μειώνεται όσο η εκπαίδευση του δικτύου εξελίσσεται, P είναι το διάνυσμα εισόδου και W τα βάρη των νευρώνων.

Το εύρος της γειτονιάς μπορεί να μεταβάλλεται κατά τη φάση εκπαίδευσης. Αρχικά η γειτονιά περιλαμβάνει όλους τους νευρώνες του επιπέδου εξόδου, όσο όμως η εκπαίδευση του δικτύου εξελίσσεται, η γειτονιά περιορίζεται σταδιακά μέχρι του σημείου να ορίζεται μόνο από ένα νευρώνα.

Ο κανόνας Kohonen επιτρέπει στα βάρη ενός νευρώνα να μάθουν ένα διάνυσμα εισόδου. Συνεπώς, ο νευρώνας που το διάνυσμα βαρών του ήταν πλησιέστερο του διανύσματος εισόδου, τροποποιείται ώστε να το προσεγγίσει ακόμη περισσότερο. Το αποτέλεσμα είναι ότι ο νευρώνας-νικητής είναι πολύ πιθανό να επικρατήσει του ανταγωνισμού την επόμενη φορά που ένα παρόμοιο διάνυσμα εισόδου παρουσιαστεί στο δίκτυο, και λιγότερο πιθανό να επικρατήσει όταν εισαχθεί ένα πολύ διαφορετικό διάνυσμα εισόδου. Όσο πιο πολλές εισοδοί παρουσιάζονται στο νευρωνικό δίκτυο, κάθε νευρώνας προσεγγίζει μια ομάδα διανυσμάτων εισόδου ταχύτερα προσαρμόζοντας τα βάρη του προς αυτά τα διανύσματα. Τελικά, εάν υπάρχουν αρκετοί νευρώνες, κάθε ομάδα παρόμοιων διανυσμάτων εισόδου θα έχει ένα νευρώνα εξόδου που δίνει ένα (1) όταν παρουσιαστεί ένα διάνυσμα που ανήκει στη συγκεκριμένη κλάση, ενώ θα δίνει μηδέν (0) σε όλες

τις άλλες περιπτώσεις. Συνεπώς, το ανταγωνιστικό δίκτυο (competitive network) μαθαίνει να κατηγοριοποιεί τα διανύσματα εισόδου που δέχεται [44].

3.8 Εφαρμογές Νευρωνικών Δικτύων

Δύο ουσιαστικά χαρακτηριστικά των Τεχνητών Νευρωνικών Δικτύων (ΤΝΔ) είναι η ικανότητα προσαρμογής τους και η μη γραμμική συμπεριφορά τους. Το πρώτο συνεπάγεται ότι ένα ΤΝΔ μπορεί να διαχειριστεί ένα σύστημα βασιζόμενο μόνο στην παρατηρούμενη συμπεριφορά του, χωρίς να απαιτείται η θεωρητική κατανόησή του. Αυτό έχει σαν αποτέλεσμα την ανάγκη ύπαρξης ενός ικανού αριθμού δειγμάτων για το πρόβλημα που καλείται να επιλύσει το ΤΝΔ.

Οι ιδιότητες αυτές σε συνδυασμό με την ικανότητά των ΤΝΔ να επιλύουν προβλήματα στα οποία τα δεδομένα τους είναι ελλιπή, ή παρουσιάζουν θόρυβο, έχει συντελέσει ώστε τα νευρωνικά δίκτυα να χρησιμοποιούνται σε ένα μεγάλο εύρος εφαρμογών, σε περιοχές τόσο διαφορετικές μεταξύ τους όπως τα οικονομικά, η ιατρική, η επιστήμη των μηχανικών, η γεωλογία και η φυσική. Στην πραγματικότητα, οπουδήποτε υπάρχουν προβλήματα πρόβλεψης, ταξινόμησης ή ελέγχου, τα ΤΝΔ μπορούν εισαχθούν και να χρησιμοποιηθούν.

Τα ΤΝΔ έχουν χρησιμοποιηθεί σε ποικίλες εφαρμογές, όπως συστήματα αυτομάτου ελέγχου, ανάλυσης δεδομένων, αναγνώρισης προτύπων (pattern recognition) και σε άλλους τομείς [33, 34, 45].

Συγκεκριμένα, τα ΤΝΔ έχουν εφαρμοστεί στη βιοτεχνολογία [46-50], στα εφαρμοσμένα μαθηματικά [51-55], σε συστήματα ανίχνευσης εισβολών [56-58], στην επεξεργασία σημάτων [59-61], στη ρομποτική [62], στη μικροηλεκτρονική [63], στην αεροδιαστημική [32], στην ηλεκτρονική [32] κ.α.

Τα νευρωνικά δίκτυα έχοντας την ικανότητα να εκτιμούν μελλοντικές καταστάσεις στηριζόμενα στην αποκτηθείσα εμπειρία από τα προϋπάρχοντα δεδομένα, μπορούν να χρησιμοποιηθούν σε προβλήματα πρόβλεψης χρονοσειρών [64-68].

Σημαντική είναι η συνεισφορά των νευρωνικών δικτύων σε προβλήματα σχετικά με **οικονομικά μεγέθη**, όπως οι τιμές των μετοχών του χρηματιστηρίου, το συνάλλαγμα, οι ομολογίες, η διεθνής κεφαλαιαγορά, οι αγορές συναλλάγματος και η εκτίμηση της τιμής του πετρελαίου, παράγοντες που καθορίζουν άμεσα την εξέλιξη των κρατών [69-73]. Συγκεκριμένα, τα ΤΝΔ χρησιμοποιούνται στην αξιολόγηση της ρευστότητας [74] και της αποδοτικότητας [75] των εταιρειών, στην εκτίμηση χρεοκοπίας εταιρειών [76, 77], στην πρόβλεψη της πορείας των χρηματοοικονομικών αγορών [78-80].

Τα ΤΝΔ έχουν βρει μεγάλη εφαρμογή στις *τραπεζικές συναλλαγές*. Χρησιμοποιούνται τόσο για την αναγνώριση της γνησιότητας των υπογραφών στις επιταγές, όσο και για την αποσαφήνιση της πιστωτικής αξιοπιστίας οργανισμών ή μεμονωμένων ατόμων που επιθυμούν να πάρουν δάνεια [32, 81]. Συγκεκριμένα, έχει αναπτυχθεί το πρόγραμμα “Νέστωρ” (Nestor), που χρησιμοποιεί νευρωνικά δίκτυα, με σκοπό τη λήψη απόφασης για την έκδοση ενός δανείου, στηριζόμενο σε προηγούμενη αποκτηθείσα γνώση. Το σύστημα αυτό έχει μεγαλύτερα ποσοστά επιτυχίας συγκριτικά με άλλες μεθόδους και είναι ιδιαίτερα χρήσιμο για αιτήσεις δανείων που υπάρχει ασάφεια ως προς τη φερεγγυότητα του δανειολήπτη [82].

Η μοντελοποίηση και η πρόβλεψη σε θέματα που αφορούν το *περιβάλλον* επιτυγχάνονται με νευρωνικά δίκτυα [83, 84]. Η δυνατότητα πρόγνωσης της νεφοκάλυψης αποτελεί σημαντική εφαρμογή των νευρωνικών δικτύων σε περιβαλλοντικά ζητήματα [85]. Επιπρόσθετα, αναφέρονται, η ανάπτυξη μοντέλων για τη διαχείριση των υδάτινων πόρων στα βουνά [86, 87], η πρόβλεψη της ποιότητας των υδάτων [88] και του αέρα [89], η ταξινόμηση των ηχογόνων πηγών [90] και η μελέτη της συγκέντρωσης του όζοντος στην Τροπόσφαιρα [91]. Επιπρόσθετα, τα νευρωνικά δίκτυα έχουν συνεισφέρει στη λήψη αποφάσεων σε αγροτικά θέματα [43, 92-97]. Η ανίχνευση της καταστροφής σε μηλιές [98] και η ταξινόμηση των κόκκων σταριού [99] με τη χρήση νευρωνικών δικτύων, είναι ορισμένες εφαρμογές των ΤΝΔ στον αγροτικό τομέα.

Τα ΤΝΔ αποτελούν μια χρήσιμη τεχνική στο τομέα της *επεξεργασίας εικόνας* [62, 100-114]. Ειδικότερα, χρησιμοποιούνται για την ταξινόμηση, την αναγνώριση αντικειμένων [108, 115-122] και τη συμπίεση μιας εικόνας [123, 124]. Σημαντική εφαρμογή είναι η αποκατάσταση μιας εικόνας και η απομάκρυνση του θορύβου [125, 126], που επιτυγχάνονται μέσω ενός αυτοσυσχετιζόμενου ΤΝΔ, το οποίο στην έξοδο παράγει την ίδια εικόνα μετά την αποκατάστασή της.

Δεν μπορεί να παραληφθεί ο ρόλος των νευρωνικών δικτύων στην *οπτική αναγνώριση* τυπωμένου ή χειρόγραφου κειμένου από τον υπολογιστή και την μετατροπή του με βάση το μέγεθος, τη γραμματοσειρά, τη θέση των χαρακτήρων και τις μουντζούρες σε ηλεκτρονικό κείμενο [127].

Τα νευρωνικά δίκτυα ενσωματώνονται σε ενσύρματα και ασύρματα *τηλεπικοινωνιακά συστήματα* [128-130]. Μια από τις πιο παλιές εφαρμογές των ΤΝΔ, που χρησιμοποιείται με επιτυχία πάνω από 30 χρόνια, είναι ένα φίλτρο με σκοπό να καταστέλλει το θόρυβο και την ηχώ από τη γραμμή, ενώ συγχρόνως περιορίζει τα σφάλματα κατά τη μετάδοση. Το πρώτο τέτοιο φίλτρο επινοήθηκε από τους Widrow και Smith, και ονομάζεται Adaline [132].

Τα τεχνητά νευρωνικά δίκτυα είναι μια ερευνητική περιοχή με έντονο ενδιαφέρον για εφαρμογές στην *Ιατρική* [32], καθώς έχουν αποδειχθεί σαν ένα αποτελεσματικό εργαλείο στην ιατρική πρόγνωση και διάγνωση ασθενειών [133-143] και άλλες ειδικότητες [144-148].

ΒΙΒΛΙΟΓΡΑΦΙΑ ΚΕΦΑΛΑΙΟΥ

1. Ritter H., Martinetz T., and Schulten K. "Neural Computation and Self-Organizing Maps - An Introduction", *Addison - Wesley* New York, 1992 (ISBN-10: 0201554429, ISBN-13: 978-0201554427).
2. Ambler P., Barrow H., Brown C., Burstall R., and Popplestone R., "A Versatile Computer-Controlled Assembly System", *Proceedings of the 3rd International Joint Conference on Artificial Intelligence*, 1973, Stanford, California, pp. 298-307.
3. Fraser H., Pugh R., Kennedy R., Ross P., and Harrison R., "A Comparison of Backpropagation and Radial Basis Functions, in the Diagnosis of Myocardial", *International Conference on Neural Networks and Expert Systems in Medicine and Healthcare*, 1994, pp. 76-84.
4. Mantzaris D., Anastassopoulos G., Tsalkidis A., and Adamopoulos A., "Intelligent Prediction of Vesicoureteral Reflux Disease", *WSEAS Transactions on Systems*, 2005, Vol. 4, Issue 9, pp. 1440-1449.
5. Mantzaris D., Anastassopoulos G., and Adamopoulos A., "Vesicoureteral Reflux Prognosis Using Artificial Neural Networks", *Proceedings of the 5th WSEAS International Conference on Simulation, Modeling and Optimization, Signal, Speech and Image Processing and Multimedia, Internet and Video Technologies*, 2005, Corfu, Greece, pp. 439-444.
6. Mantzaris D., Anastassopoulos G., Adamopoulos A., and Gardikis S., "A Non-Symbolic Implementation of Abdominal Pain Estimation in Childhood", *Information Science*, 2008, Vol. 178, Issue 20, pp. 3860-3866.
7. Mantzaris D., Anastassopoulos G., Adamopoulos A., Stephanakis I., Kambouri K., and Gardikis S., "Abdominal Pain Estimation in Childhood based on Artificial Neural Network Classification" *Proceedings of the 10th International Conference on Engineering Applications of Neural Networks*, 2007, Thessaloniki Hellas pp. 129-134.
8. Mantzaris D., Anastassopoulos C., and Lymperopoulos K., "Medical Disease Prediction Using Artificial Neural Networks", *Proceedings of 8th IEEE International Conference on BioInformatics and BioEngineering*, 2008, Athens Greece.
9. Mantzaris D., Anastassopoulos G., Iliadis L., and Adamopoulos A., "An Evolutionary Technique for Medical Diagnostic Risk Factors Selection", *Proceedings of the 5th IFIP Conference on Artificial Intelligence Applications & Innovations (ALAI'2009)*, accepted.
10. Mantzaris D., Anastassopoulos G., Adamopoulos A., Stephanakis I., Kambouri K., and Gardikis S., "Selective Clinical Estimation of Childhood Abdominal Pain Based on Pruned Artificial Neural Networks" *Proceedings of the 3rd WSEAS International*

- Conference on Cellular and Molecular Biology, Biophysics and Bioengineering (BIO'07), Athens, Greece, 2007, pp. 50-55.*
11. Kurzweil R., “The Age of Intelligent Machines” *MIT Press* Cambridge, 1990.
 12. Βλαχάβας Ι., Κεφαλάς Π., Βασιλειάδης Ν., Κόκκορας Φ., Σακελλαρίου Η., “Τεχνητή Νοημοσύνη”, Γ' Έκδοση, *Εκδόσεις Β.Γκιούρδας Εκδοτική*, ISBN: 960-387-431-0, 2006.
 13. Haykin S., “Neural Networks: A Comprehensive Foundation”, *Prentice Hall*, PTR Upper Saddle River NJ USA, 1998 (ISBN-10: 0132733501, ISBN-13: 9780132733502).
 14. Leischner A., “Paul Broca and Significance of his Works for Clinical Pathology of the Brain”, *Bratislavské Lekárske Listy*, 1972, Vol. 57, No. 5, pp. 615–623.
 15. McCulloch W., and Pitts W., “A Logical Calculus of Ideas Immanent in Nervous Activity”, *Bulletin of Mathematical Biophysics*, 1943, Vol. 5, pp. 115-133.
 16. Hebb D., “The Organization of Behavior”, *Wiley-Interscience*, New York, 1949.
 17. Bryson A. and Ho Y., “Applied Optimal Control: Optimization, Estimation, and Control”, *Ginn and Company*, Waltham, MA, 1969.
 18. Minsky M., “Theory of Neural-Analog Reinforcement Systems and its Application to the Brain-Model Problem”, *PhD thesis, Princeton University*, 1954.
 19. Minsky M., “Steps toward Artificial Intelligence”, *Proceedings of the Institute of Radio Engineers (IRE)*, 1961, Vol. 49, Issue 1, pp. 8-30.
 20. McCarthy J., Minsky M.L., Rochester N., and Shannon C.E., “A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence”, August 31, 1955.
 21. Rochester N., Holland H., Habit H., and Duda L., “Tests on a cell assembly theory of the action of the brain, using a large digital computer” *IRE Transactions on Information Theory*, IT-2 1956 pp. 80-93.
 22. Rosenblatt F., “The perceptron: A Probabilistic Model for Information Storage and Organization in the Brain” *Psychological Review*, 1958, Vol. 65, No. 6, pp. 386-408.
 23. Widrow B., and Hoff M., “Adaptive Switching Circuits”, *WESCON Convention, Record Part 4*, New York, 1960 pp. 96-104.
 24. Minsky M., and Papert S., “Perceptrons”, *MA: MIT Press*, Cambridge, 1969.
 25. Kohonen T., “Correlation Matrix Memories”, *IEEE Transactions on Computers*, 1972, Vol. 21, pp. 353-359.
 26. Anderson J., “A Simple Neural Network Generating an Interactive Memory”, *Mathematical Biosciences*, 1972, Vol. 14, pp. 197-220.

27. Grossberg S., "Adaptive Pattern Classification and Universal Recoding I: Parallel Development and Coding of Neural Feature Detectors", *Biological Cybernetics*, 1976, Vol. 23, pp. 121-134.
28. Λυκοθανάσης Σ., "Γενετικοί Αλγόριθμοι κι Εφαρμογές" *Ελληνικό Ανοικτό Πανεπιστήμιο, Σχολή Θετικών Επιστημών και Τεχνολογίας Πάτρα*, 2001.
29. Hopfield J., "Neural Networks and Physical Systems with Emergent Collective Computational Abilities", *Proceedings of the National Academy of Sciences*, 1982, Vol. 79, pp. 2554-2558.
30. Rumelhart D., and McClelland J., "Parallel Distributed Processing: Explorations in the Microstructure of Cognition", *MIT Press: Cambridge Mass*, 1986.
31. Rall W., "Perspectives on Neuronal Modelling", "The Segmental Motor System," *Oxford University Press*, NY, 1990, pp. 129-149.
32. Demuth H., Beale M., and Hagan M., "Neural Network Toolbox™ User's Guide" *MathWorks, Inc* 2008.
33. Nayak R., Jain L., and Ting B., "Artificial Neural Networks in Biomedical Engineering: A Review" *Proceedings of the 1st Asian-Pacific Conference on Advance Computation*, 2001, Sydney Australia, pp. 887-892.
34. Dayhoff J., and DeLeo J., "Artificial Neural Networks Opening the Black Box" *CANCER Supplement*, 2001, Vol. 91, No. 8, pp. 1615-1635.
35. Astion M., Wener M., Thomas R., Hunder G., and Bloch D., "Overtraining in Neural Networks that Interpret Clinical Data", *Clinical Chemistry*, 1993, Vol. 39, pp. 1998-2004.
36. Akcayol M., Cinar C., Bulbul H., and Kilicarslan A., "Artificial Neural Network Based Modeling of Injection Pressure in Diesel Engines", *WSEAS Transactions on Computers*, 2004, Vol. 3, Issue 5, pp. 1538-1544.
37. Coury D., and Segatto E., "Pattern Recognition to Distinguish Magnetizing Inrush from Internal Faults in Power Transformers", *WSEAS Transactions on Systems*, 2004, Vol. 3, Issue 3, pp. 1258-1264.
38. Zhang G., and Berardi V., "An Investigation of Neural Networks in Thyroid Function Diagnosis", *Health Care Management Science*, 1998, Vol. 1, No. 1, pp. 29-37.
39. Parzen E., "On Estimation of a Probability Density Function and Mode", *Annals of Mathematical Statistics*, 1962, Vol. 33, No.3, pp. 1065-1076.
40. Anderson D., and McNeil G., "Artificial Neural Networks Technology", *Kaman Science Corporation*, 1992.

41. Dieterlea F., Müller-Hagedorn S., Liebich H., and Gauglitz G., “Urinary Nucleosides as Potential Tumor Markers Evaluated by Learning Vector Quantization”, *Artificial Intelligence in Medicine*, 2003, Vol. 28, No. 3, pp. 265-279.
42. Bilbao J., and Bravo E., “Comparing Newton with ANN for Load Flow” *WSEAS Transactions on Mathematics*, 2004, Vol. 3, Issue 1, pp. 31-36.
43. Diamantopoulou M., “Comparative Analysis of Different Artificial Neural Network Approaches for Fir Trees Breast Height Form Factor Estimator”, *Proceedings of the 10th International Conference on Engineering Applications of Neural Networks*, 2007, Thessaloniki, Greece, pp. 352-359.
44. Konar A., “Artificial Intelligence and Soft Computing: Behavioral Cognitive Modeling of the Human Brain”, *CRC Press LLC*, 2000.
45. Κινγ Ρ., “Υπολογιστική Νοημοσύνη στον Έλεγχο Συστημάτων”, *Τραυλός ΙΙ*, Αθήνα, 1998.
46. Hanai T., and Honda H., “Application of Knowledge Information Processing Methods to Biochemical Engineering, Biomedical and Bioinformatics Fields”, *Adv. Biochemical Engineering Biotechnology*, 2004, Vol. 91, pp. 51-73.
47. Krogh A., “What are artificial neural Networks?”, *National Biotechnology*, 2008, Vol. 26, No. 2, pp. 195-197.
48. Montague G., and Morris J., “Neural-network Contributions in Biotechnology”, *Trends Biotechnology*, 1994, Vol. 12, No. 8, pp. 312-324.
49. Schetin V., and Brazhnikov A., “Diagnostic Rule Extraction Using Neural Networks”, *Biomedical Engineering*, 2000, Vol. 34, No. 1, pp. 17-23.
50. Zhu Y, Williams S, and Zwiggelhaar R. “Computer Technology in Detection and Staging of Prostate Carcinoma: A Review”. *Medical Image Analysis*, 2006, Vol. 10, No. 2, pp. 178-199.
51. Huang D., “A Constructive Approach for Finding Arbitrary Roots of Polynomials by Neural Networks”, *IEEE Transactions on Neural Networks*, 2004, Vol. 15, No. 2, pp. 477 - 491.
52. Kinderman L., Lewandowski A., and Protzel P., “A framework for Solving Functional Equations with Neural Networks”, *Proceedings of the 8th International Conference on Neural Information Processing (ICONIP)*, 2001, Vol. 2, pp. 1075-1078.
53. Margaritis A., and Adamopoulos M., “Solving Nonlinear Algebraic Systems Using Artificial Neural Networks”, *Proceedings of the 10th International Conference on Engineering Applications of Neural Networks*, 2007, Thessaloniki, Greece, pp. 107-120.

54. Perantonis S., Ampazis N., Varoufakis S., and Antoniou G., "Constrained Learning in Neural Networks: Application to Stable Factorization of 2-D Polynomials" *Neural Processing Letters*, 1998 Vol. 7, No. 1, pp. 5-14.
55. Rao M., and Ravindran K., "An Evolutionary Approach for Solving Differential Equations", *WSEAS Transactions on Circuits*, 2002, Vol. 1, Issue 1, pp. 100 - 105.
56. Zhang Z., and Manikopoulos C., "Comparative Evaluation of Classification Algorithms in Anomaly Intrusion Detection", *WSEAS Transactions on Computer*, 2004, Vol. 3, Issue 5, pp. 1551-1556.
57. Zhang Z., Li J., Manicopoulos C., Jorgenson J., and Ucles J., "Neural Networks in Statistical Anomaly Intrusion Detection", *Proceedings of the 5th World Multiconference on Circuits, Systems, Communications & Computers* 2001.
58. Zhang Z., Li J., Manicopoulos C., Jorgenson J., and Ucles J., "HIDE: a Hierarchical Neural Intrusion Detection System Using Statistical Preprocessing and Neural Network Classification" *Proceedings of the 2001 IEEE Workshop on Information Assurance and Security*, N.Y., 2001.
59. Campos M., and Carpenter G., "S-TREE: Self-Organizing Trees for Data Clustering and Online Vector Quantization", *Neural Networks*, 2001, Vol. 14, Issues 4-5, pp. 505-525.
60. Luo F.-L., and Unbehauen R., "Applied Neural Networks for Signal Processing", *New York: Cambridge University Press, Cambridge UK*, 1997.
61. Zaknick A., "Introduction to the Modified Probabilistic Neural Network for General Signal Processing", *IEEE Transactions on Signal Processing*, 1998, Vol. 46, No.7, pp. 1980-1990.
62. Carnimeo L., "Target Recognition in Mobile Robot Vision Systems via Neural Networks with Local Interconnections", *Proceedings of the 3rd WSEAS International Conference on Neural Network and Applications*, 2002, pp. 107-111.
63. Varshavsky V., Marakhovsky V., and Levin I., "Artificial Neurons Based on CMOS β -Driven Threshold Elements with Functional Inputs", *WSEAS Transactions on Systems*, 2004, Vol. 3, Issue 2, pp. 442 - 448.
64. Back A., and Tsoi A., "FIR and IIR Synapses, a New Neural Network Architecture for Times Series Modelling", *Neural Computation*, 1991, Vol. 3, No.3, pp. 375-385.
65. Li H., and Kozma R., "A Dynamic Neural Network Method for Time Series Prediction Using the KIII Model", *Proceedings of International Joint Conference on Neural Networks*, 2003, pp. 347-353.

66. Priel A., and Kanter I., "Time Series Generation by Recurrent Neural Networks", *Annals of Mathematics and Artificial Intelligence*, 2003, Vol. 39, No. 3, pp. 315-332.
67. Small M., and Tse C., "Minimum Description Length Neural Networks for Time Series Prediction", *Physical Review*, 2002, Vol. 66, No. 6, pp. 066701-066701 -066701-066712.
68. Zaknich A., De Silva C., and Attikiouzel Y., "The Probabilistic Neural Network for Nonlinear Time Series Analysis", *Proceedings of the IEEE International Joint Conference on Neural Networks*, Singapore, 1991.
69. Amin-Naseri M., and Gharacheh E., "A Hybrid Artificial Intelligence Approach to Monthly Forecasting of Crude Oil Price Time Series", *Proceedings of the 10th International Conference on Engineering Applications of Neural Networks*, 2007, Thessaloniki, Greece, pp. 160-167.
70. Azoff E., "Neural Network Time Series Forecasting of Financial Markets", *John Wiley & Sons Chichester*, England, 1994.
71. Papadourakis G., Spanoudakis G., and Gotsiass A., "Application of Neural Networks in Short-Term Stock Price Forecasting", *Proceedings of the 1st International Workshop Neural Networks in the Capital Markets* 1993.
72. Zapranis A., and Alexandridis A., "Weather Derivatives Pricing: Modeling the Seasonal Residual Variance of an Ornstein-Uhlenbeck Temperature Process with Neural Networks", *Proceedings of the 10th International Conference on Engineering Applications of Neural Networks*, 2007, Thessaloniki, Greece, pp. 178-190.
73. Hussain A., Knowles A., Lisboa P., and El-Deredey W., "Financial Time Series Prediction Using Polynomial Pipelined Neural Networks", *Expert Systems with Applications: An International Journal*, 2008, Vol. 35, Issue 3, pp. 1186-1199
74. Chen N., and Liu Y., "Application of Artificial Neural Network to Corporate Liquidity Determination", *The 3rd International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, 2007, pp. 557-560.
75. Loukeris N., "Radial Basis Function Networks to Hybrid Neuro-Genetic - RBFNS in Financial Evaluation of Corporations", *International Journal of Computers*, 2008, Vol. 2, pp. 176-182.
76. Anandarajan M., Lee P., and Anandarajan A., "Bankruptcy Prediction of Financially Stressed Firms: An Examination of the Predictive Accuracy of Artificial Neural Networks", *Intelligent Systems in Accounting, Finance and Management*, 2001, Vol. 10, pp. 69-81.
77. Atiya A., "Bankruptcy Prediction for Credit Risk Using Neural Networks: A Survey and New Results", *IEEE Transactions on Neural Networks*, 2001, Vol. 12, pp. 929-935.

78. Biscontri R., “A Radial Basis Function Approach to Earnings Forecast”, *Intelligent Systems in Accounting, Finance and Management*, 2012, Vol. 19, pp. 1-18.
79. Liang L., and Wu D., “An Application of Pattern Recognition on Scoring Chinese Corporations Financial Conditions Based in Backpropagation Neural Network”, *Computers & Operations Research*, 2005, Vol. 32, pp. 1115-1129.
80. Mostafa M., “Modeling the Competitive Market Efficiency of Egyptian Companies: A Probabilistic Neural Network Analysis”, *Expert Systems with Applications*, 2009, Vol. 39, pp. 8839-8838.
81. Hajek P., and Olej V., “Municipal Creditworthiness Modeling by Clustering Methods”, *Proceedings of the 10th International Conference on Engineering Applications of Neural Networks*, 2007, Thessaloniki, Greece, pp. 168-177.
82. McNelis P., “Neural Networks in Finance: Gaining Predictive Edge in the Market”, *Elsevier Academic Press*, USA, ISBN: 0-12-485967-4, 2005.
83. Aitkenhead M., Mustard M., and McDonald A., “Using Neural Networks to Predict Spatial Structure in Ecological Systems”, *Ecological Modelling*, 2004, Vol. 179, No. 3, pp. 393-403.
84. Dowla F., and Rogers L., “Solving Problems in Environmental Engineering and Geosciences with Artificial Neural Networks”, *Massachusetts Institute of Technology*, USA, 1995.
85. Mantzari V., and Mantzaris D., “Solar Radiation: Cloudiness Forecasting Using a Soft Computing Approach”, *Artificial Intelligence Research*, 2013, Vol. 2, No. 1, pp. 69-80.
86. Glezakos T., Tsiligridis T., Iliadis L., Yialouris C., Maris F., and Ferentinos K., “Feature Extraction for Time Series Data: an Artificial Neural Network Evolutionary Training Models for the Management of Mountainous Watersheds”, *Proceedings of the 10th International Conference on Engineering Applications of Neural Networks*, 2007, Thessaloniki, Greece, pp. 410-418.
87. Iliadis L., and Maris F., “An Artificial Neural Network Model for Mountainous Water-Resources Management. The Case of a Cyprus Mountainous Watersheds”, *Environmental Modeling and Software*, Elsevier, 2007, Vol. 22, Issue 7, pp. 1066-1072.
88. Partalas I., Hatzikos E., Tsoumakas G., and Vlahavas I., “Ensemble Selection of Water Quality Prediction”, *Proceedings of the 10th International Conference on Engineering Applications of Neural Networks*, 2007, Thessaloniki, Greece, pp. 428-435.
89. Karatzas K., and Kaltsatos S., “Modeling and Forecasting of Air Quality with the Aid of Machine Learning Methods in Thessaloniki, Greece”, *Proceedings of the 10th*

- International Conference on Engineering Applications of Neural Networks*, 2007, Thessaloniki, Greece, pp. 419-427.
90. Stoeckle S., Pah N., Kumar D., and MacLachlan N., “Environmental Sound Sources Classification by Using Neural Networks”, *Proceedings of the Seventh Australian and New Zealand Intelligent Systems Conference*, 2001, pp. 399-404.
91. Paschalidou A., Iliadis L., Kassomenos P., and Bezirtzoglou C., “Neural Modelling of the Tropospheric Ozone Concentrations in an Urban Site”, *Proceedings of the 10th International Conference on Engineering Applications of Neural Networks*, 2007, Thessaloniki, Greece, pp. 436-445.
92. Abainech B., Nayak R., Ranjan A., and Gomes J., “On-line Control of Fed-Batch Fermentation by Using SOM Based Multiple Local Linear Models”, *Proceedings of the 10th International Conference on Engineering Applications of Neural Networks*, 2007, Thessaloniki, Greece, pp. 368-375.
93. Diamantopoulou M., Georgiou P., and Papamichail D., “Runoff Modeling through Cascade Correlation and Back Propagation Artificial Neural Networks”, *Proceedings of the 10th International Conference on Engineering Applications of Neural Networks*, 2007, Thessaloniki, Greece, pp. 360-367.
94. Diamantopoulou M., “Artificial Neural Networks as an Alternative Tool in Pine Bark Volume Estimation”, *Computer and Electronics Agriculture*, 2005, Vol. 48, No. 3, pp. 235-244.
95. Diamantopoulou M., “Tree-bole Volume Estimation on Standing Pine Trees Using Cascade Artificial Neural Network Models”, *Agriculture Engineering International-CIGR Journal of Scientific Research and Development*, 2006, Vol. VIII, pp. 1-14.
96. Diamantopoulou M., “Predicting Fir Trees Stem Diameters Using Artificial Neural Network Models”, *Southern African Forestry Journal*, 2005, Vol. 205, No. 1, pp. 39-44.
97. Corne S., Carver S., Kunin W., Lennon J., and Van Hees W., “Predicting Forest Attributes in Southeast Alaska Using Artificial Neural Networks”, *Forest Science*, 2004, Vol. 50, No. 2, pp. 259-279.
98. Kwater T., Bartaman J., and Pekala R., “Artificial Neural Networks in Detection of Apples Damage”, *Proceedings of the 10th International Conference on Engineering Applications of Neural Networks*, 2007, Thessaloniki, Greece, pp. 376-379.
99. Paliwal J., Visen N., and Jayas D., “Evaluation of Neural Network Architectures for Cereal Grain Classification Using Morphological Features”, *Journal of Agricultural Engineering Research*, 2001, Vol. 79, No. 4, pp. 361-370.

100. Ahmed M., and Farag A., “Two-Stage Neural Network for Volume Segmentation of Medical Images”, *Pattern Recognition Letters*, 1997, Vol. 18, Issues 11-13, pp. 1143-1151.
101. Aizenberg I., Aizenberg N., Hiltner J., Moraga C., and Meyer zu Bexten E., “Cellular Neural Networks and Computational Intelligence in Medical Image Processing”, *Image and Vision Computing*, 2001, Vol. 19, No. 4, pp. 177-183.
102. Gomolka Z., “Homogenous Neural Network Prepared for Interferometry Images”, *IMACS Symposium on Mathematical Modelling*, 2000, Vienna, pp. 555-558.
103. Hall L., Bensaid A., Clarke L., Velthuizen R., Silbiger M., and Bezdek J., “A Comparison of Neural Network and Fuzzy Clustering Techniques in Segmenting Magnetic Resonance Images of the Brain”, *IEEE Transactions on Neural Networks*, 1992, Vol. 3, No. 5, pp. 672-682.
104. Ossen A., Zamzow T., Oswald H., and Fleck E., “Segmentation of Medical Images Using Neural Network Classifiers”, *International Conference on Neural Networks and Expert Systems in Medicine and Healthcare*, 1994, pp. 427-432.
105. Özkan M., Sprenkels H., and Dawant B., “MultiSpectral Magnetic Resonance Image Segmentation Using Neural Networks”, *International Joint Conference on Neural Networks*, 1990, Vol. 1, pp. 429-434.
106. Carpenter G., Grossberg S., and Mehanian C., “Invariant Recognition of Cluttered Scenes by a Self-Organizing ART Architecture: CORT-X Boundary Segmentation”, *Neural Networks*, 1989, Vol. 2, No. 3, pp. 169-181.
107. Egmont-Petersena M., de Ridderb D., and Handels H., “Image Processing with Neural Networks – A Review”, *Pattern Recognition*, 2002, Vol. 35, pp. 2279–2301.
108. Garcia C., and Delakis M. “Convolutional Face Finder: A Neural Architecture for Fast and Robust Face Detection”, *IEEE Transactions on Pattern Analysis and Machine Intelligenc*, 2004, Vol. 26, No. 11, pp. 1408-1423.
109. Hisashi A., Tsuyoshi O., Takahashi N., and Tanaka M., “Sigma-Delta Cellular Neural Network for 2D Modulation”, *Neural Networks*, 2008, Vol. 21, Issue 2-3, pp. 349-357.
110. Li J., Manry M., Narasimha P., and Yu C., “Feature Selection Using a Piecewise Linear Network”, *IEEE Transactions on Neural Network*, 2006, Vol. 17, No. 5, pp. 1101 – 1115.
111. Narasimha P., Manry M., and MaldonadocF., “Upper Bound on Pattern Storage in Feedforward Networks”, *Neurocomputing (Letters)*, 2008, Vol. 71, pp. 3612– 3616.

112. Narasimha P., Delashmit W., Manry M., Li J., and Maldonado F., “An Integrated Growing-Pruning Method for Feedforward Network Training”, *Neurocomputing*, 2008, Vol. 71, pp. 2831–2847.
113. Narasimha P., Manry M., and Maldonado F., “Upper Bound on Pattern Storage in Feedforward Networks”, *Proceedings of International Joint Conference on Neural Networks*, 2007, Orlando, Florida, USA.
114. Zhao Z.Q., Huang D.S.A., and Sun B.Y., “Human Face Recognition Based on Multi-Features Using Neural Networks Committee”, *Physical Review Letters*, 2004, Vol. 25, No. 12, pp. 1351-1358.
115. Hitam M., Muslan Y., Deris M., and Saman M., “Image Texture Classification Using Projection and Neural Network”, *WSEAS Transactions on Systems*, 2004, Vol. 3, No. 4, pp. 1825-1830.
116. Kuhn R., Bordas R., Wunderlich B., Michaelis B., and Thevenin D., “Colour Class Identification of Tracers Using Artificial Neural Networks”, *Proceedings of the 10th International Conference on Engineering Applications of Neural Networks*, 2007, Thessaloniki, Greece, pp. 387-394.
117. Matsumoto H., Masumoto R., and Kuroda C., “Application Methods for Self Organizing Map in Process Imaging for Dynamic Behavior of Aerated Agitation Vessel” *Proceedings of the 10th International Conference on Engineering Applications of Neural Networks*, 2007, Thessaloniki, Greece, pp. 210-220.
118. Ohta E., Mitsukura Y., Fukumi M., Akamatsu N., and Yasutomo M., “An Extraction Method of Liver Tumors by Using Genetic Algorithms and Neural Networks”, *Proceedings of the Seventh Australian and New Zealand Intelligent Systems Conference*, 2001, pp. 63-66.
119. O’Leary T., Mikel U., and Becker R., “Computer-Assisted Image Interpretation: Use of a Neural Network to Differentiate Tubular Carcinoma From Sclerosing Adenosis”, *Modern Pathology*, 1992, Vol. 5, No. 4, pp. 402-405.
120. Phung S., Chai D., and Bouzerdoum A., “Skin Colour Based Face Detection”, *Proceedings of the Seventh Australian and New Zealand Intelligent Systems Conference*, 2001, pp. 171-176.
121. Piuri V., and Scotti F., “Morphological Classification of Blood Leucocytes by Microscope Images”, *IEEE International Conference on Computational Intelligence for Measurement Systems and Applications*, 2004, pp. 103-108.

122. Rowley H., Baluja S., and Kanade T., "Neural Network-Based Face Detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998, Vol. 20, No. 1, pp. 23-38.
123. Giusti D., Eduardo A., and L. L., "Images Compression for Medical Diagnosis Using Neural Networks", *Journal of Computer Sciences and Technology*, 2000, Vol. 1, No.1, pp.76 - 81.
124. Gomolka Z., Twarog B., and Kwiatkoski B., "Data Compression in Biometrics Images", *Proceedings of the 10th International Conference on Engineering Applications of Neural Networks*, 2007 Thessaloniki, Greece, pp. 403-409.
125. Gendy S., Kothapalli G., and Bouzerdoum A., "A Fast Algorithm for Image Restoration Using a Recurrent Neural Network with Bound-Constrained Quadratic Optimization", *Proceedings of the Seventh Australian and New Zealand Intelligent Systems Conference*, 2001, pp. 111-116.
126. Anthony D., Hines E., Barham J., and Taylor D., "A Comparison of Image Compression by Neural Networks and Principal Component Analysis", *International Joint Conference on Neural Networks*, 1990, Vol. 1, pp. 339-344.
127. Lee S., "Off-Line Recognition of Totally Unconstrained Handwritten Numerals Using Multilayer Cluster Neural Network", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1996, Vol. 18, No. 6, pp. 646-652.
128. Haralambous H., Papadopoulos H., and Economou L., "Using Neural Networks for Predicting Likelihood of Interference to Groundwave Users in the HF Spectrum", *Proceedings of the 10th International Conference on Engineering Applications of Neural Networks*, 2007, Thessaloniki, Greece, pp. 200-209.
129. Murillo-Fuentes J., and González-Serrano F., "Applying GCMAC to Predistortion in GSM Base Stations", *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'99)*, 1999, Vol. 5, pp. 2575-2578.
130. Young J., Hanselmann T., Zaknich A., and Attikiouzel Y., "Adaptive Complex Modified Probabilistic Neural Network in Digital Channel Equalization", *Proceedings of the Seventh Australian and New Zealand Intelligent Systems Conference*, 2001, pp. 247-252.
131. Αργυράκης Π., "Νευρωνικά Δίκτυα και Εφαρμογές", *Ελληνικό Ανοικτό Πανεπιστήμιο*, Πάτρα Ελλάδα, 2000.
132. Keogan M., Lo J., Freed K., Raptopoulos V., Blake S., Kamel I., Weisinger K., Rosen M., and Nelson R., "Outcome Analysis of Patients with Acute Pancreatitis by Using an Artificial Neural Network", *Academic Radiology* 2002, Vol. 9, No. 4, pp. 410-419.

133. Dayhoff J., and DeLeo J., “Artificial Neural Networks Opening the Black Box” *CANCER Supplement*, 2001, Vol. 91, No. 8, pp. 1615-1635.
134. Ishak W., and Siraj F., “Artificial Intelligence in Medical Application: an Exploration”, *Health Informatics Europe Journal [Online]*, 2002, (Online article at: <http://www.hi-europe.info/files/2002/9980.htm>).
135. Claude R., Charles-Daniel A., Jean A., and Jean-Francois G., “Bibliometric Overview of the Utilization of Artificial Neural Networks in Medicine and Biology”, *Scientometrics*, 2004, Vol. 59, No. 1, pp. 117-130.
136. Economou G.-P., Lymberopoulos D., and Goutis C., “An ANNs-Based System for the Diagnosis and Treatment of Diseases”, *Journal of Neural Processing Letters*, 1995, Vol. 2, No. 1, pp. 22-26.
137. Jankowski N., “Approximation and Classification in Medicine with IncNet Neural Networks”, *Workshop on Machine Learning and Applications: Machine Learning in Medical Applications*, 1999, Greece, pp. 53-58.
138. Lisboa P., “A Review of Evidence of Health Benefit from Artificial Neural Networks in Medical Intervention”, *Neural Networks*, 2002, Vol. 15, Issue 1, pp.11-39.
139. Larder B., Wang D., and Revell A., “Application of Artificial Neural Networks for Decision Support in Medicine”, *Journal of Medical Systems*, 2008, Vol. 32, No. 5, pp. 403-408.
140. Maiellaro P., Cozzolongo R., and Marino P., “Artificial Neural Networks for the Prediction of Response to Interferon Plus Ribavirin Treatment in Patients with Chronic Hepatitis C”, *Current Pharmaceutical Design*, 2004, Vol. 10, pp. 2101-2109.
141. Patel J., and Goyal R., “Applications of Artificial Neural Networks In Medical Science”, *Current Clinical Pharmacology*, 2007, Vol. 2, No. 3, pp. 217-226.
142. Salim N., “Medical Diagnosis Using Neural Networks”, *Faculty of Information Technology, University Utara Malaysia*, 2004, Sintok, Kedah.
143. Sordo M., “Introduction to Neural Networks in Healthcare”, *Open Clinical Website*, 2002.
144. Eftekhari B., Mohammad K., Ardebili H., Ghodsi M., and Ketabchi E., “Comparison of Artificial Neural Network and Logistic Regression Models for Prediction of Mortality in Head Trauma Based on Initial Clinical Data”, *BMC Medical Informatics and Decision Making*, 2005, Vol. 5, No. 3, pp. 1-8.
145. Garliauskas A., “Self-Organization of Mosaics in Artificial Neural Networks for the Visual Cortex of the Brain”, *Proceedings of the Seventh Australian and New Zealand Intelligent Systems Conference*, 2001, pp. 259-264.

146. Lopes C., Ludermir T., de Souto M., and Ludermir A., “Neural Networks for the Analysis of Common Mental Disorders Factors”, *Proceedings of the VII Brazilian Symposium on Neural Networks (SBRN'02)*, 2002, pp. 114.
147. Monoyiou E., Ventouras E., Ktonas P., Paparrigopoulos T., Dikeos D., Uzunoglou N., and Soldatos C., “Multi-Layer Perceptrons for the Detection of Sleep EEG Transient Waveforms”, *Proceedings of the Fourth International Conference on Neural Networks and Expert Systems in Medicine and Healthcare*, 2001, Milos Island Greece, pp. 47-50.
148. Vourkas M., Papadourakis G., Ktonas P., and Micheloyannis S., “Effects of Data Segmentation on the Discrimination of Three Mental States Using ANN and Different EEG Signal Representations”, *Proc. of the Fourth International Conference on Neural Networks and Expert Systems in Medicine and Healthcare*, 2001, Milos Island Greece, pp. 65-68.

ΕΥΦΥΕΙΣ ΥΠΟΛΟΓΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ ΣΤΗΝ ΑΞΙΟΛΟΓΗΣΗ ΤΗΣ ΑΠΟΔΟΤΙΚΟΤΗΤΑΣ ΕΙΣΗΓΜΕΝΩΝ ΕΤΑΙΡΕΙΩΝ



4.1 Εισαγωγή

Μέχρι τώρα δόθηκαν τα κυριότερα σημεία για τις τεχνικές χρηματοοικονομικής αξιολόγησης της αποδοτικότητας των εταιρειών. Ιδιαίτερη ανάλυση υπήρξε για τη μέθοδο των χρηματοοικονομικών αριθμοδεικτών, η οποία χρησιμοποιείται ευρύτατα στην αξιολόγηση επιχειρήσεων από οικονομικής πλευράς. Επιπρόσθετα, παρουσιάστηκαν τα Τεχνητά Νευρωνικά Δίκτυα, οι αρχιτεκτονικές, οι βασικές τους λειτουργίες, οι αλγόριθμοι εκπαίδευσης κι οι συναρτήσεις μεταφοράς τους. Το επόμενο στάδιο είναι η πρακτική εφαρμογή των νευρωνικών δικτύων για την εκτίμηση της αποδοτικότητας ενός συνόλου εταιρειών εισηγμένων στο Χρηματιστήριο Αξιών Αθηνών. Συγκεκριμένα, αναπτύχθηκαν ευφυείς αλγόριθμοι για την αξιολόγηση της αποδοτικότητας των προαναφερθέντων εταιρειών. Οι υλοποιήσεις και τα αποτελέσματα που προέκυψαν αναπτύσσονται στο παρόν κεφάλαιο. Η προσέγγιση αυτή αποτελεί την πρώτη υπολογιστική ευφυή μέθοδο για την αξιολόγηση της αποδοτικότητας των ελληνικών εταιρειών του Χρηματιστηρίου Αξιών Αθηνών.

Ο συνδυασμός των διαθέσιμων χρηματοοικονομικών αριθμοδεικτών για την αξιολόγηση της αποδοτικότητας των επιχειρήσεων, όπως αυτά προτάθηκαν από τους ειδικούς οικονομολόγους, και της γνώσης των μηχανικών σε νευρωνικά δίκτυα, οδήγησαν στην ανάπτυξη ενός ΤΝΔ για την ταξινόμηση των συγκεκριμένων εταιρειών ανάλογα με την αποδοτικότητά τους. Η προτεινόμενη αρχιτεκτονική και τα αποτελέσματα που προέκυψαν παρουσιάζονται στη συνέχεια.

4.2 Συλλογή Δεδομένων

Η παρούσα διατριβή βασίστηκε σε ένα τυχαίο δείγμα εταιρειών εισηγμένων στο Χρηματιστήριο Αξιών Αθηνών (Χ.Α.Α.) για την περίοδο 2006 έως και 2009. Το δείγμα αποτελείται από 30 εταιρείες διαφόρων κλάδων του Χ.Α.Α για τις οποίες υπολογίστηκαν εννέα χρηματοοικονομικοί δείκτες που αξιολογούν την αποδοτικότητα των συγκεκριμένων εταιρειών. Η

επιλογή των αριθμοδεικτών βασίστηκε στη δημοτικότητά τους στη διεθνή βιβλιογραφία [1-6], καθώς και στο θεωρητικό τους υπόβαθρο.

Συγκεκριμένα, οι χρηματοοικονομικοί δείκτες που χρησιμοποιήθηκαν είναι:

1. Ο δείκτης γενικής ρευστότητας (R1),
2. Ο δείκτης ειδικής ρευστότητας (R2),
3. Ο δείκτης ίδια προς ξένα κεφάλαια (R3),
4. Ο δείκτης αποδοτικότητας ιδίων προς αποδοτικότητα συνολικών κεφαλαίων (R4),
5. Ο δείκτης χρηματοοικονομικής μόχλευσης (R5),
6. Ο δείκτης ταχύτητας κυκλοφορίας αποθεμάτων (R6),
7. Ο δείκτης ταχύτητας κυκλοφορίας απαιτήσεων (R7),
8. Ο δείκτης δανειακής επιβάρυνσης (R8) και
9. Ο δείκτης ταχύτητας κυκλοφορίας βραχυπρόθεσμων υποχρεώσεων (R9).

Ο Πινάκας 4.1 παρουσιάζει τους προαναφερθέντες αριθμοδείκτες (1^η στήλη) και τους μαθηματικούς τύπους αυτών στη 2^η στήλη του Πίνακα.

Η **γενική ρευστότητα** αποτελεί έναν από τους βασικότερους αριθμοδείκτες της ρευστότητας και δείχνει το πλήθος των φορών που το κυκλοφορούν ενεργητικό καλύπτει τις βραχυπρόθεσμες υποχρεώσεις. Όταν ο δείκτης αυτός όταν λαμβάνει τιμή μεγαλύτερη της μονάδας παρέχεται η ένδειξη ότι η επιχείρηση διαχειρίζεται αποτελεσματικά το κεφάλαιο κίνησης. Παρ' όλα αυτά η τελική εκτίμηση για την αξιολόγηση της ρευστότητας προϋποθέτει την περαιτέρω διερεύνηση με τη χρήση αριθμοδεικτών δραστηριότητας όπου αξιολογείται ο βαθμός εντατικής χρήσης των κυκλοφορούντων περιουσιακών στοιχείων.

Για τον υπολογισμό της **ειδικής ρευστότητας** λαμβάνονται υπόψη τα πιο άμεσα ρευστοποιήσιμα στοιχεία, που είναι οι απαιτήσεις και τα διαθέσιμα. Ένας υψηλός αριθμοδείκτης ειδικής ρευστότητας παρέχει την ένδειξη ότι η επιχείρηση έχει τη δυναμική να καλύψει τις βραχυπρόθεσμες υποχρεώσεις της από τα αμυντικά περιουσιακά στοιχεία της. Παρόλα αυτά και εδώ η αξιολόγηση του δείκτη προϋποθέτει τη χρήση αριθμοδεικτών δραστηριότητας και ειδικότερα τους αριθμοδείκτες ταχύτητας κυκλοφορίας απαιτήσεων και ταχύτητα κυκλοφορίας βραχυπρόθεσμων υποχρεώσεων.

Ο αριθμοδείκτης **ίδια προς ξένα κεφάλαια** είναι ιδιαίτερα σημαντικός δεδομένου ότι δείχνει το βαθμό κάλυψης των ξένων κεφαλαίων από ίδια. Μια υψηλή τιμή του συγκεκριμένου αριθμοδείκτη δημιουργεί "αίσθημα" ασφαλείας στους πιστωτές της επιχείρησης δεδομένου ότι σε σημαντικό βαθμό η χρηματοδότησή της γίνεται από ίδια κεφάλαια. Αυτό μπορεί να συνεπάγεται μεγαλύτερα όρια πιστώσεων αλλά και μικρότερο χρηματοοικονομικό κόστος. Παρόλα αυτά ένας υψηλός αριθμοδείκτης ιδίων προς ξένα μπορεί να συνδέεται με χαμηλή απόδοση των ιδίων κεφαλαίων.

Αριθμοδείκτες	Μαθηματικός Τύπος
Δείκτης γενικής ρευστότητας	$\frac{\text{Κυκλοφορούν ενεργητικό}}{\text{βραχυπρόθεσμες υποχρεώσεις}}$
Δείκτης ειδικής ρευστότητας	$\frac{\text{κυκλοφορούν ενεργητικό} - \text{αποθέματα}}{\text{βραχυπρόθεσμες υποχρεώσεις}}$
Δείκτης ίδια προς ξένα κεφάλαια,	$\frac{\text{ίδια κεφάλαια}}{\text{Ξένα κεφάλαια}}$
Δείκτης αποδοτικότητας ιδίων προς αποδοτικότητα συνολικών κεφαλαίων	$\frac{\text{R. O. E.}}{\text{R. O. A.}}$
Δείκτης χρηματοοικονομικής μόχλευσης	$\frac{\text{Ενεργητικό}}{\text{Ίδια Κεφάλαια}}$
Δείκτης ταχύτητας κυκλοφορίας αποθεμάτων	$\frac{\text{Κόστος πωληθέντων}}{\text{Μέσο ύψος αποθεμάτων}}$
Δείκτης ταχύτητας κυκλοφορίας απαιτήσεων	$\frac{\text{Πωλήσεις}}{\text{Μέσο ύψος απαιτήσεων}}$
Δείκτης δανειακής επιβάρυνσης	$\frac{\text{Ξένα κεφάλαια}}{\text{Συνολικά κεφάλαια}}$
Δείκτης ταχύτητας κυκλοφορίας βραχυπρόθεσμων υποχρεώσεων.	$\frac{\text{Αγορές}}{\text{Μέσο ύψος βραχυπρόθεσμων υποχρεώσεων}}$

Πίνακας 4.1. Αριθμοδείκτες και τύπος υπολογισμού τους.

Ο αριθμοδείκτης *αποδοτικότητα; ιδίων προς αποδοτικότητα συνολικών κεφαλαίων (ROE/ROA)* αξιολογεί την ύπαρξη ή μη πλεονεκτήματος μόχλευσης. Όταν ο δείκτης αυτός λαμβάνει τιμή μεγαλύτερη της μονάδας λαμβάνεται η ένδειξη ότι η επιχείρηση ευνοείται από τη χρήση των ξένων κεφαλαίων. Με άλλα λόγια ο δανεισμός έχει θετική επίδραση στα αποτελέσματα της επιχείρησης. Προκειμένου, όμως, να υπολογιστεί το επίπεδο μόχλευσης της

επιχείρησης χρησιμοποιείται ο *δείκτης χρηματοοικονομικής μόχλευσης* του παρόντος υποδείγματος. Ο δείκτης αυτός δείχνει το βαθμό κάλυψης των ιδίων κεφαλαίων από το σύνολο του ενεργητικού.

Οι δυο *δείκτες ταχύτητας κυκλοφορίας αποθεμάτων* και *απαιτήσεων* αποτελούν όπως έχει αναφερθεί, συμπληρωματικούς δείκτες των βασικών δεικτών ρευστότητας, δεδομένου ότι αξιολογούν το βαθμό εντατικής χρήσης των κυκλοφορούντων στοιχείων. Στο σημείο αυτό απαιτείται η αναφορά στο περιεχόμενο του *βαθμού εξυπηρέτησης των βραχυπρόθεσμων υποχρεώσεων*.

Τέλος, ο *δείκτης δανειακής επιβάρυνσης* παρέχει εικόνα για την επιλογή κεφαλαιακής διάρθρωσης μιας επιχείρησης. Το πόσο επιτυχής είναι αυτή η επιλογή αξιολογείται ως ένα βαθμό από τους δυο δείκτες που συνδέονται με τη μόχλευση και που έχουν οριστεί σε προγενέστερη παράγραφο.

Από τα ανωτέρω, γίνεται αντιληπτό ότι η ρευστότητα, η κεφαλαιακή διάρθρωση και η δραστηριότητα αποτελούν τρία χρηματοοικονομικά πεδία τα οποία σε συνδυασμό παρέχουν κρίσιμα συμπεράσματα για τη γενικότερη απόδοση της επιχείρησης.

Ο Πίνακας 4.2 αποτυπώνει τους συντελεστές συσχέτισης των χρησιμοποιούμενων δεικτών. Ο σκοπός ενός ΤΝΔ είναι να εκτελεί μια συγκεκριμένη λειτουργία, συνδυάζοντας ένα σύνολο εισόδων με τις επιθυμητές εξόδους, αναγκάζοντας το ΤΝΔ να εκτελέσει τους απαραίτητους υπολογισμούς για την τροποποίηση των συναπτικών βαρών μεταξύ των τεχνητών νευρώνων. Στη διαδικασία αυτή δεν υπάρχει κανένας περιορισμός στη φάση εκπαίδευσης καθώς και στη φάση προσομοίωσης [7, 8].

<i>Χρηματοοικονομικοί Δείκτες</i>	<i>R1</i>	<i>R2</i>	<i>R3</i>	<i>R4</i>	<i>R5</i>	<i>R6</i>	<i>R7</i>	<i>R8</i>	<i>R9</i>
R1	1.00								
R2	0.81	1.00							
R3	0.19	0.26	1.00						
R4	-0.12	-0.07	-0.04	1.00					
R5	-0.36	-0.41	-0.20	0.25	1.00				
R6	-0.13	0.10	-0.05	0.16	-0.08	1.00			
R7	-0.05	-0.21	-0.05	0.04	-0.15	-0.10	1.00		
R8	-0.51	0.63	-0.41	0.18	0.69	0.05	-0.02	1.00	
R9	0.64	0.56	0.11	-0.10	-0.41	0.09	0.36	-0.45	1.00

Πίνακας 4.2. Συντελεστές συσχέτισης των εννέα χρηματοοικονομικών δεικτών.

Ο Πίνακας 4.3 συνοψίζει τα περιγραφικά στατιστικά των χρηματοοικονομικών δεικτών που χρησιμοποιήθηκαν στην αξιολόγηση των εταιρειών για την περίοδο 2006 έως και 2009.

<i>Στατιστικά Μεγέθη</i>	<i>R1</i>	<i>R2</i>	<i>R3</i>	<i>R4</i>	<i>R5</i>	<i>R6</i>	<i>R7</i>	<i>R8</i>	<i>R9</i>
Μέσος όρος	1.57	0.99	1.17	1.43	3.00	5.31	2.56	0.59	1.35
Διάμεσος	1.34	0.79	0.56	0.96	2.67	3.77	2.29	0.62	1.19
Τυπική απόκλιση	1.03	0.69	3.68	5.06	1.88	5.23	1.41	0.16	0.79
Διακύμανση	1.06	0.48	13.58	25.56	3.54	27.40	1.98	0.03	0.63
Ελάχιστη τιμή	0.34	0.30	0.04	-14.30	0.33	0.35	0.60	0.14	0.00
Μέγιστη τιμή	7.70	4.38	40.00	33.30	15.07	26.80	6.85	0.96	5.70
Εύρος τιμών	7.36	4.08	39.96	47.60	14.74	26.45	6.25	0.82	5.70
1 ^ο Τεταρτημόριο	0.97	0.53	0.38	0.34	1.99	1.49	1.69	0.50	0.84
3 ^ο Τεταρτημόριο	1.80	1.10	0.99	1.79	3.37	7.99	2.96	0.62	1.19

Πίνακας 4.3. Τιμές περιγραφικών στατιστικών των χρηματοοικονομικών δεικτών.

Οι εταιρείες χαρακτηρίστηκαν σαν αποδοτικές και μη-αποδοτικές ανάλογα με τις τιμές των εννέα χρηματοοικονομικών δεκτών. Ο Πίνακας 4.4 παρουσιάζει τις αξιολογήσεις των 30 εταιρειών ανά έτος.

<i>Επίπεδα αποδοτικότητας</i>	<i>Έτη</i>			
	<i>2006</i>	<i>2007</i>	<i>2008</i>	<i>2009</i>
Αποδοτική	15	14	8	11
Μη-αποδοτική	15	16	22	19

Πίνακας 4.4. Ταξινόμηση των εταιρειών σύμφωνα με τη χρηματοοικονομική ανάλυση.

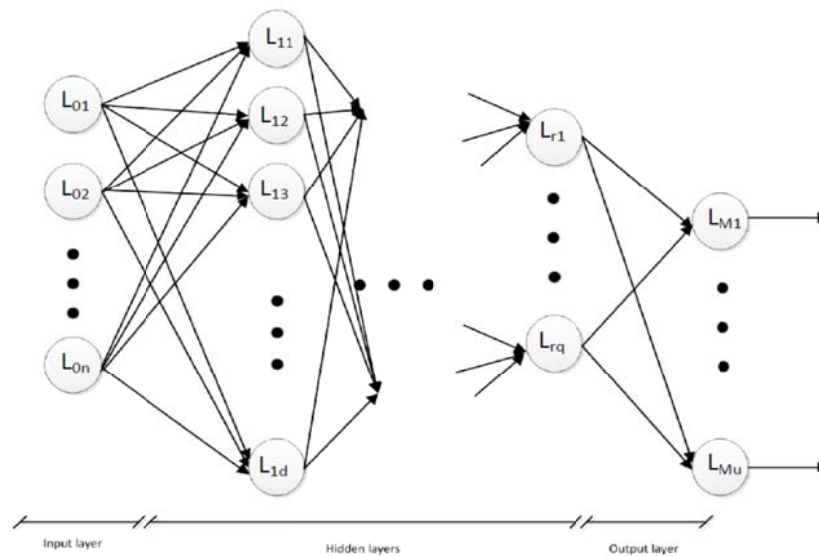
Η μελέτη αυτή χρησιμοποίησε ένα σύνολο δεδομένων αποτελούμενο από 120 δείγματα, εκ των οποίων τα 48 (40%) ήταν αποδοτικές εταιρείες και τα υπόλοιπα 72 (60%) μη-αποδοτικές. Το

συγκεκριμένο σύνολο δεδομένων χωρίστηκε σε δυο υποσύνολα, ένα αποτελούμενο από 70 καταχωρήσεις για την εκπαίδευση (training set) του MLP και την κατασκευή του PNN και ένα άλλο από 50 δείγματα για τον έλεγχο των υλοποιημένων ΤΝΔ (testing set). Το σύνολο εκπαίδευσης περιέχει 28 αποδοτικές εταιρείες και 42 μη-αποδοτικές. Το σύνολο ελέγχου περιλαμβάνει 20 αποδοτικές και 30 μη-αποδοτικές εταιρείες.

4.3 Νευρωνικά Δίκτυα για την Ευφυή Αξιολόγηση της Εταιρικής Αποδοτικότητας

Τα πολύ-στρωματικά νευρωνικά δίκτυα (Multi-Layer Perceptrons – MLP Neural Networks) είναι δίκτυα εμπρόσθιας διάδοσης δεδομένων που χρησιμοποιούν αλγόριθμους οπισθοδιάδοσης σφάλματος για την εκπαίδευσή τους [9, 10]. Η συσχέτιση των εξαρτημένων και ανεξάρτητων μεταβλητών αποτελεί ένα σημαντικό χαρακτηριστικό για τα MLPs, με αποτέλεσμα να μπορούν να χρησιμοποιηθούν στην επεξεργασία δεδομένων οικονομικής φύσης.

Ένα MLP αποτελείται από ένα στρώμα εισόδου και ένα στρώμα εξόδου. Στο στρώμα εισόδου, ο αριθμός των κόμβων εισόδου ισούται με τον αριθμό των μεταβλητών του προβλήματος, ενώ ο αριθμός των νευρώνων στο στρώμα εξόδου καθορίζεται από τις απαιτήσεις του προβλήματος [11]. Οι νευρώνες των στρωμάτων μεταξύ του επιπέδου εισόδου και του επιπέδου εξόδου είναι τα κρυμμένα στρώματα των νευρωνικών δικτύων, με αποτέλεσμα ένα MLP νευρωνικό δίκτυο να έχει M επίπεδα συναπτικών συνδέσεων και $M+1$ στρώματα νευρώνων. Τα κρυμμένα στρώματα και ένας μεγάλος αριθμός χαρακτηριστικών της αρχιτεκτονικής MLP προσδιορίζονται με τη μέθοδο δοκιμής και σφάλματος. Το Σχήμα 4.1 απεικονίζει την δομή ενός MLP νευρωνικού δικτύου. Κάθε νευρώνας αντιπροσωπεύεται με το συμβολισμό L_{mk} , όπου m είναι ο αριθμός του στρώματος του δικτύου και k είναι ο νευρώνας του στρώματος m . Το m κυμαίνεται από 0 έως M , ενώ τα βάρη και οι πολώσεις (biases) έχουν εύρος μεταξύ m και $m-1$. Τα επίπεδα του παριστάνονται με τους συμβολισμούς W_m και b_m , αντίστοιχα. Το W_m είναι ένας $n_m \times n_{m-1}$ πίνακας, όπου n_m , και n_{m-1} είναι οι αριθμοί των νευρώνων των m και $m-1$ στρωμάτων, αντίστοιχα.



Σχήμα 4.1. Αρχιτεκτονική πολύ-στρωματικού δικτύου.

Το διάνυσμα εξόδου L_M για ένα συγκεκριμένο διάνυσμα εισόδου L_0 υπολογίζεται από την ακόλουθη σχέση:

$$L_{mk} = f \left(\sum_{j=1}^{n_{m-1}} W_{kj}^m L_{(m-1)j} + b_k^m \right) \quad (4.1)$$

όπου $f(\cdot)$ είναι η συνάρτηση μεταφοράς των νευρώνων του m επιπέδου

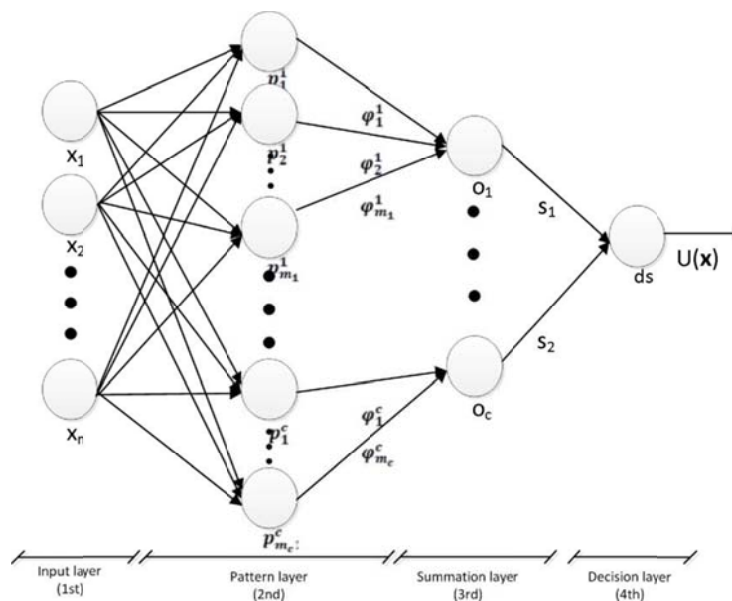
Παρόλο που τα MLPs έχουν χρησιμοποιηθεί σε ένα ευρύ φάσμα εφαρμογών, οι ερευνητές εξακολουθούν να εμφανίζονται επιφυλακτικοί. Ο λόγος για αυτή την αντιπαράθεση προέρχεται από το χαρακτηριστικό του “μαύρου κουτιού” των MLPs, καθώς μπορούν να ανιχνεύσουν κρυφές συσχετίσεις στα δεδομένα. Σε αντίθεση με το ευρετικό χαρακτηριστικό των MLPs, τα PNNs που προσεγγίζουν Bayesian στατιστικές τεχνικές έχουν την ικανότητα να ταξινομήν διανύσματα με μια διαδικασία η οποία είναι παρόμοια με τον τρόπο που ο άνθρωπος λαμβάνει τις αποφάσεις του [12].

Τα PNNs προτάθηκαν και αναπτύχθηκαν από τον Specht (1990). Ο Parzen υλοποιεί την συνάρτηση πυκνότητας πιθανότητας (Probabilistic Density Function - PDF) (Probabilistic Density Function – PDF) σαν εκτιμητή στις εμπρόσθια διάδοσης αρχιτεκτονικές νευρωνικών δικτύων [13]. Έστω τα x_1, x_2, \dots, x_n που αναπαριστούν μια ακολουθία δειγμάτων μιας μονοδιάστατης, στατιστικά ανεξάρτητης τυχαίας μεταβλητής με άγνωστη συνάρτηση πιθανοκρατικής πυκνότητας $f(x)$, όπου μπορεί να εκτιμηθεί από το δείγμα, ως εξής:

$$f(x) = \frac{1}{n} \left(\sum_{i=1}^n \frac{1}{\lambda_n} K \left(\frac{x - x_i}{\lambda_n} \right) \right) \quad (4.2)$$

όπου $K(\cdot)$ είναι ένας ουσιαστικός εκτιμητής στο κέντρο κάθε δείγματος και λ είναι ένας παράγοντας εξομάλυνσης που εξαρτάται από το μέγεθος του δείγματος, n .

Τα PNNs είναι κατάλληλα για προβλήματα που η έξοδος τους έχει πολλές κλάσεις στοχεύοντας στο συνδυασμό νέων διανυσμάτων εισόδων με τα αποθηκευμένα δεδομένα στα δίκτυα, προκειμένου να ταξινομήσουν σωστά τα δεδομένα εισόδου σε μία από τις διαθέσιμες κατηγορίες στόχων του προβλήματος. Η συγκεκριμένη διαδικασία είναι παρόμοια με την ανθρώπινη συμπεριφορά [12, 14]. Ένα PNN είναι τεσσάρων επιπέδων εμπρόσθιας διάδοσης νευρωνικό δίκτυο, αποτελούμενο από ένα στρώμα εισόδου, ένα επίπεδο προτύπων (pattern layer), ένα στρώμα άθροισης (summation layer) και ένα επίπεδο λήψης απόφασης (decision-making layer) [15, 16], και απεικονίζεται στο Σχήμα 4.2. Το πρώτο επίπεδο μεταφέρει τα δεδομένα εισόδου στους νευρώνες του επιπέδου προτύπων. Υπάρχει ένα νευρώνα στο στρώμα εισόδου για κάθε μεταβλητή πρόβλεψης. Το στρώμα προτύπων υπολογίζει τις αποστάσεις μεταξύ του διανύσματος εισόδου και των διανυσμάτων εκπαίδευσης και παράγει ένα διάνυσμα, τα στοιχεία του οποίου δείχνουν πόσο κοντά σε ένα δείγμα εκπαίδευσης μπορεί να είναι το διάνυσμα εισόδου. Αυτό το στρώμα έχει ένα νευρώνα για κάθε στοιχείο του συνόλου δεδομένων εκπαίδευσης.



Σχήμα 4.2. Αρχιτεκτονική πιθανοκρατικού νευρωνικού δικτύου.

Στην περίπτωση που ένα πρότυπο x μεταφέρεται από το στρώμα εισόδου στο στρώμα προτύπων, ο νευρώνας p_j του στρώματος προτύπων υπολογίζει την έξοδό του με την ακόλουθη εξίσωση:

$$\varphi_j^i(x) = \frac{1}{(2\pi)^{\frac{d}{2}} \sigma^d} \exp \left[-\frac{(x - p_j^i)^T (x - p_j^i)}{2\sigma^2} \right] \quad (4.3)$$

όπου το d συμβολίζει τη διάσταση του προτύπου x , σ είναι η παράμετρος εξομάλυνσης της ακτινικής συνάρτησης βάσης και p_j ο νευρώνας του διανύσματος.

Οι νευρώνες του στρώματος άθροισης (O_i) υπολογίζουν τη μέγιστη πιθανότητα του προτύπου x για να ταξινομηθεί στην κλάση u_i . Αυτό επιτυγχάνεται με τον υπολογισμό του αθροίσματος και του μέσου όρου των νευρώνων της εξόδου που ανήκουν στην ίδια κλάση, σύμφωνα με την ακόλουθη εξίσωση:

$$s_i(x) = \frac{1}{(2\pi)^{\frac{d}{2}} \sigma^d} \frac{1}{m_i} \sum_{j=1}^{m_i} \exp \left[-\frac{(x - p_j^i)^T (x - p_j^i)}{2\sigma^2} \right] \quad (4.4)$$

όπου το m_i δηλώνει το συνολικό αριθμό των δειγμάτων στην κατηγορία U_i . Στην περίπτωση που οι πιθανότητες, καθώς και οι εσφαλμένες κατατάξεις της κάθε κατηγορίας είναι εκ των προτέρων ίδιες, ο νευρώνας του στρώματος λήψης απόφασης ταξινομεί το πρότυπο x σύμφωνα με τον κανόνα απόφαση του Baye με βάση των αποτελεσμάτων όλων των νευρώνων του στρώματος άθροισης. $C(s)$ είναι η εκτιμώμενη τάξη του προτύπου x και c είναι ο συνολικός αριθμός των κλάσεων κατά την κατασκευή του PNN [17, 18].

Τα PNNs αποτελούν ελκυστική αρχιτεκτονική νευρωνικού δικτύου, λόγω των χαρακτηριστικών τους [19]. Ειδικότερα, τα PNNs δεν απαιτούν προϋποθέσεις σχετικά με την κατανομή των τυχαίων μεταβλητών που χρησιμοποιήθηκαν για την εκτίμηση, και μπορούν να χειριστούν ακόμα και κατανομές πολλαπλών μοντέλων. Επιπλέον, τα PNNs μπορούν να μάθουν από τα πρότυπα εκπαίδευσης με ένα μόνο πέρασμα, χωρίς την αναγκαιότητα επαναληπτικής εκπαίδευσης. Τα υπάρχοντα βάρη δεν χρειάζεται να τροποποιηθούν, αλλά είναι νέα διανύσματα προσαρτώνται στους πίνακες βαρών κατά τη διάρκεια της φάσης κατασκευής του συγκεκριμένου PNN. Αυτό το χαρακτηριστικό σημαίνει ότι τα PNNs μπορούν να διαχειρίζονται μεγάλες

ποσότητες δεδομένων ταχύτερα από ότι τα MLP νευρωνικά δίκτυα και μπορούν να χρησιμοποιηθούν σε εφαρμογές πραγματικού χρόνου [20, 21].

Τα νευρωνικά δίκτυα παρέχουν, γενικά, βελτιωμένη απόδοση με κανονικοποίηση των δεδομένων. Η χρήση των πρωτότυπων δεδομένων σαν είσοδο του νευρωνικού δικτύου μπορεί να προκαλέσει πρόβλημα σύγκλισης. Συνεπώς, τα σύνολα δεδομένων εκπαίδευσης και δοκιμής μετατρέπονται σε τιμές μεταξύ -1 και 1, με την ακόλουθη εξίσωση:

$$nv = \frac{\text{min_value} - \text{act_value}}{\text{max_value} - \text{min_value}} \quad (4.5)$$

όπου max_value και min_value είναι οι μέγιστες και ελάχιστες τιμές, act_value είναι η πραγματική τιμή και nv είναι η κανονικοποιημένη τιμή.

Ο κύριος στόχος της κανονικοποίησης, σε συνδυασμό με την αρχικοποίηση των βαρών, είναι να επιτραπεί η λειτουργία της συνάρτησης συμπίεσης τουλάχιστον στην αρχή της φάσης εκμάθησης. Συνεπώς, η κλίση, η οποία είναι συνάρτηση της παραγώγου, θα λαμβάνει πάντοτε τιμές διαφορετικές από το μηδέν. Στο τέλος του κάθε αλγορίθμου, οι έξοδοι μετατρέπονται στην αρχική μορφή των δεδομένων προκειμένου να επιτευχθεί το επιθυμητό αποτέλεσμα.

Ο αριθμός των νευρώνων του επιπέδου εισόδου καθορίζεται από τις παραμέτρους του προβλήματος. Στο εδάφιο 4.2 διευκρινίστηκε ότι η αξιολόγηση της αποδοτικότητας των εταιρειών βασίζεται σε εννέα (9) χρηματοοικονομικούς δείκτες. Κατά συνέπεια, στην παρούσα μελέτη, το στρώμα εισόδου των ΤΝΔ που έχουν υλοποιηθεί αποτελείται από εννέα (9) νευρώνες. Οι παράμετροι εισόδου των ΤΝΔ, καθώς και η κωδικοποίησή τους, παρουσιάζονται στην 1^η στήλη του Πίνακα 4.5 και αντιστοιχούν στις εισόδους του νευρωνικού δικτύου με τους χρηματοοικονομικούς αριθμοδείκτες οι οποίοι παρουσιάζονται στη 2^η στήλη. Η 3^η στήλη απεικονίζει το εύρος κανονικοποίησης για κάθε μεταβλητή.

<i>Είσοδοι TNA</i>	<i>Παράμετροι</i>	<i>Κανονικοποίηση</i>
1	Δείκτης γενικής ρευστότητας (R1)	[-1,1]
2	Δείκτης ειδικής ρευστότητας (R2)	[-1,1]
3	Δείκτης ίδια προς ξένα κεφάλαια (R3)	[-1,1]
4	Δείκτης αποδοτικότητας ιδίων προς αποδοτικότητα συνολικών κεφαλαίων (R4)	[-1,1]
5	Δείκτης χρηματοοικονομικής μόχλευσης (R5)	[-1,1]
6	Δείκτης ταχύτητας κυκλοφορίας αποθεμάτων (R6)	[-1,1]
7	Δείκτης ταχύτητας κυκλοφορίας απαιτήσεων (R7)	[-1,1]
8	Δείκτης δανειακής επιβάρυνσης (R8)	[-1,1]
9	Δείκτης ταχύτητας κυκλοφορίας βραχυπρόθεσμων υποχρεώσεων (R9)	[-1,1]

Πίνακας 4.5. Κωδικοποίηση των παραμέτρων της αποδοτικότητας των εταιρειών.

Η αποτελεσματικότητα των εταιρειών κατατάσσεται σε 2 επίπεδα, σύμφωνα με την ανάλυση στο εδάφιο 4.2. Οι βαθμίδες αυτές απαιτείται να κωδικοποιηθούν προκειμένου να χρησιμοποιηθούν από το νευρωνικό δίκτυο. Ο Πίνακας 4.6 παρουσιάζει την κωδικοποίηση του κάθε επιπέδου αποτελεσματικότητας των εταιρειών. Η 1^η στήλη του Πίνακα παρουσιάζει περιγραφικά το επίπεδο αποτελεσματικότητας και η 2^η στήλη την κωδικοποίησή του, όπως χρησιμοποιήθηκε από το νευρωνικό δίκτυο.

<i>Επίπεδο Αποδοτικότητας</i>	<i>Κωδικοποίηση</i>
Μη αποδοτικό	0
Αποδοτικό	1

Πίνακας 4.6. Επίπεδο αποδοτικότητας και κωδικοποίηση

Ο αριθμός των νευρώνων του επιπέδου εξόδου ορίζεται τόσο από τον επιθυμητό αριθμό των μεταβλητών του προβλήματος όσο και τον τύπο του TNA. Συγκεκριμένα, το MLP απαιτεί ένα

νευρώνα στο στρώμα εξόδου για τον καθορισμό του επιπέδου αποτελεσματικότητας της εταιρείας. Από την άλλη πλευρά, η δομή του PNN δικτύου απαιτεί τόσους νευρώνες, όσες είναι οι πιθανές εκβάσεις. Στη συγκεκριμένη εφαρμογή υπάρχουν δυο (2) επίπεδα αποτελεσματικότητας, επομένως, υπάρχουν 2 νευρώνες στο επίπεδο εξόδου του PNN.

Οι χρησιμοποιούμενες συναρτήσεις μεταφοράς για την αρχιτεκτονική MLP ήταν η hyperbolic tangent sigmoid για το κρυφό επίπεδο και η γραμμική (linear) συνάρτηση για το επίπεδο εξόδου [22, 23]. Η ακτινική συνάρτηση βάσης (radial basis function) και η ανταγωνιστική (competitive) συνάρτηση εφαρμόστηκαν για το κρυφό επίπεδο και το επίπεδο εξόδου, αντίστοιχα των αρχιτεκτονικών PNN.

Ο προσδιορισμός του αριθμού των νευρώνων στο κρυφό επίπεδο του MLP επιτεύχθηκε με τη διαδικασία δοκιμής και σφάλματος. Μια υπολογιστική μέθοδος υλοποιήθηκε με σκοπό την τροποποίηση του αριθμού των νευρώνων στο κρυμμένο στρώμα και τον υπολογισμό της απόδοσης όλων των τοπολογιών ΤΝΔ. Ο αλγόριθμος Levenberg-Marquardt επιλέχθηκε για την εκπαίδευση των MLPs, καθώς είναι ένας εύρωστος αλγόριθμος, κατάλληλος για τα μη-γραμμικά ελαχίστου-τετραγώνου προβλήματα [23].

Σε αντίθεση με τα MLPs όπου ο αριθμός των κρυφών στρωμάτων δεν είναι πλήρως καθορισμένος, η δομή των PNNs έχει μόνο ένα κρυφό στρώμα. Επιπλέον, ο αριθμός των νευρώνων στο κρυφό στρώμα των PNN εξαρτάται από τον αριθμό των προτύπων κατά τη διάρκεια της κατασκευής τους. Ο αριθμός των κόμβων για το κρυφό στρώμα PNN είναι ο αριθμός των προτύπων κατά τη διάρκεια της κατασκευής του. Κατά συνέπεια, το προτεινόμενο PNN είχε 70 νευρώνες για το κρυφό επίπεδο, καθώς τα διαθέσιμα στοιχεία για την κατασκευή του PNN, αποτελούνταν από 70 περιπτώσεις. Η σχεδίαση των PNNs είναι απλή και δεν βασίζεται σε εκπαίδευση, με αποτέλεσμα να μην επιλεγεί αλγόριθμος εκμάθησης κατά τη διάρκεια υλοποίησης του PNN για το συγκεκριμένο πρόβλημα [23].

Το μέσο τετραγωνικό σφάλμα (Mean Squared Error – MSE) [23] χρησιμοποιήθηκε σαν κριτήριο για την αξιολόγηση της απόδοσης των MLPs, του οποίου η μαθηματική εξίσωση είναι:

$$\text{MSE} = \frac{1}{N} \sum_{K=1}^N e^2(k) = \frac{1}{N} \sum_{K=1}^N [t(k) - a(k)]^2 \quad (4.6)$$

όπου N είναι ο αριθμός των προτύπων, t(k), a(k) και e(k) είναι η επιθυμητή τιμή, η υπολογιζόμενη από το MLP νευρωνικό δίκτυο τιμή και η τιμή σφάλματος για το k πρότυπο, αντίστοιχα.

Ο προσδιορισμός της αρχιτεκτονικής PNN για την αξιολόγηση της αποδοτικότητας των εταιρειών βασίζεται στα διαθέσιμα χαρακτηριστικά του συγκεκριμένου προβλήματος, καθώς και

στον καθορισμό της “παραμέτρου εξομάλυνσης” (“smoothing parameter”) ή παράγοντα εξάπλωσης (spread factor), που αντιπροσωπεύει το πλάτος της υπολογιζόμενης Gaussian καμπύλης για κάθε συνάρτηση πυκνότητας πιθανότητας. Στην παρούσα μελέτη, ο παράγοντας εξάπλωσης κυμαίνεται από 0.1 έως 50, με βήμα 0.1 [20, 24].

4.4 Αποτελέσματα Προσομοίωσης

Η εκπαίδευση και προσομοίωση των διάφορων αρχιτεκτονικών για την αξιολόγηση της αποδοτικότητας των εισηγμένων στο Χρηματιστήριο Αξιών Αθηνών πραγματοποιήθηκε με το Neural Network Toolbox που ενσωματώνει το MATLAB, Mathworks Inc [23]. Το πρόγραμμα αυτό επιλέχθηκε λόγω της ευχρηστίας, της αποτελεσματικότητάς και του φιλικού περιβάλλοντος διασύνδεσης χρήστη.

Τα χαρακτηριστικά των MLP νευρωνικών δικτύων που υλοποιήθηκαν παρουσιάζονται στον Πίνακα 4.7. Η 2^η στήλη περιγράφει τον αριθμό νευρώνων κάθε επιπέδου του MLP. Τα δίκτυα που υλοποιήθηκαν περιείχαν, πέρα από τα επίπεδα εισόδου και εξόδου, ένα κρυφό επίπεδο όπου μεταβάλλονταν ο αριθμός των νευρώνων του κρυφού επιπέδου ή το επιθυμητό MSE ή και οι δυο προαναφερθείσες παράμετροι. Η 3^η στήλη αποτυπώνει τις συναρτήσεις μεταφοράς, οι οποίες ήταν η tansig και η purelin. Ο αλγόριθμος Levenberg-Marquardt χρησιμοποιήθηκε για την εκπαίδευση των MLPs και παρουσιάζεται στην 4^η στήλη του Πίνακα. Η 5^η στήλη καταγράφει το επιθυμητό MSE, όπως ορίστηκε κατά τη φάση εκπαίδευσης. Στον Πίνακα 4.7 παρουσιάζονται νευρωνικά δίκτυα με όλα τα τεχνικά χαρακτηριστικά τους να είναι ίδια. Αυτό οφείλεται στο γεγονός ότι κατά τη φάση εκπαίδευσης, τα αρχικά βάρη και οι πολώσεις δεν είναι τα ίδια για όλα τα MLPs, αλλά ποικίλουν, διατηρώντας αναλλοίωτες τις άλλες παραμέτρους. Όμως, οι αρχικές τιμές των βαρών και των πολώσεων δεν μπορούν να εμφανιστούν στον πίνακα. Οι διαφορετικές αρχικές τιμές των βαρών και των πολώσεων δημιουργούν νευρωνικά δίκτυα με διαφορετικές επιδόσεις, καθώς τα τελικά βάρη και οι πολώσεις δεν έχουν τις ίδιες τιμές σε όλα τα νευρωνικά δίκτυα με αμετάβλητα τα υπόλοιπα χαρακτηριστικά. Ενδεικτικά, το πρώτο και δεύτερο μοντέλο νευρωνικού δικτύου του Πίνακα 4.7. έχουν την ίδια αρχιτεκτονική, παρόλο που οι αρχικές συνθήκες ήταν διαφορετικές. Παρόμοιες ρυθμίσεις εφαρμόστηκαν στις περιπτώσεις των 4-7 και 8-10 μοντέλων νευρωνικών δικτύων, προκειμένου να εκπαιδευτούν και να κατασκευαστούν ΤΝΔ για την αξιολόγηση της αποδοτικότητας των εταιρειών. Είναι προφανές ότι διαφορετικές αρχικές συνθήκες για την εκπαίδευση των MLPs συνεπάγονται μεταβολή της απόδοσης των νευρωνικών δικτύων.

<i>No.</i>	<i>Αρχιτεκτονική ΤΝΔ</i>	<i>Συνάρτηση Μεταφοράς</i>	<i>Αλγόριθμος Εκπαίδευσης</i>	<i>Επιθυμητό MSE</i>
1	9-4-1	tansig purelin	Levenberg-Marquardt	0.5
2	9-4-1	tansig purelin	Levenberg-Marquardt	0.5
3	9-4-1	tansig purelin	Levenberg-Marquardt	0.8
4	9-6-1	tansig purelin	Levenberg-Marquardt	0.5
5	9-6-1	tansig purelin	Levenberg-Marquardt	0.5
6	9-6-1	tansig purelin	Levenberg-Marquardt	0.5
7	9-6-1	tansig purelin	Levenberg-Marquardt	0.5
8	9-6-1	tansig purelin	Levenberg-Marquardt	0.8
9	9-6-1	tansig purelin	Levenberg-Marquardt	0.9
10	9-6-1	tansig purelin	Levenberg-Marquardt	0.9
11	9-6-1	compet purelin	Levenberg-Marquardt	0.5
12	9-9-1	tansig purelin	Levenberg-Marquardt	0.5

Πίνακας 4.7. Χαρακτηριστικά αρχιτεκτονικών MLP για την αξιολόγηση της αποδοτικότητας των εταιρειών.

Ο Πίνακας 4.8 παρουσιάζει τα αποτελέσματα που προέκυψαν από την εφαρμογή των MLPs στην αξιολόγηση της αποδοτικότητας των εισηγμένων στο Χρηματιστήριο Αξιών Αθηνών εταιρειών. Η 2^η στήλη προβάλλει τις αρχιτεκτονικές των MLP νευρωνικών δικτύων. Τα ποσοστά επιτυχούς ταξινόμησης για τα δεδομένα ελέγχου, εκπαίδευσης και τα συνολικά δίνονται στην 3^η, 4^η και 5^η στήλη, αντίστοιχα. Με άλλα λόγια, η 3^η, 4^η και 5^η στήλη απεικονίζουν τα σωστά ταξινομημένα δεδομένα επί των 50, 70 και 120 διανυσμάτων, αντίστοιχα. Η 6^η, 7^η και 8^η στήλη

παρουσιάζουν τα ποσοστά επιτυχούς ταξινόμησης για τις εταιρείες ανεπαρκούς αποτελεσματικότητας για τα δεδομένα ελέγχου, εκπαίδευσης και τα συνολικά δεδομένα, αντίστοιχα.

No.	Αρχιτεκτονική ΤΝΔ	Ποσοστά Επιτυχούς Ταξινόμησης			Ποσοστά Επιτυχούς Ταξινόμησης		
		Σύνολο Ελέγχου	Σύνολο Εκπαίδευσης	Σύνολο Δεδομένων	Σύνολο Ελέγχου	Σύνολο Εκπαίδευσης	Σύνολο Δεδομένων
1	9-4-1	28.21	28.72	28.32	33.80	34.91	34.21
2	9-4-1	31.99	31.92	31.06	33.80	37.42	37.47
3	9-4-1	31.99	33.03	32.18	38.53	39.91	38.65
4	9-6-1	61.99	63.03	62.55	67.71	67.98	67.47
5	9-6-1	61.99	63.03	62.55	67.18	69.91	67.47
6	9-6-1	74.25	76.30	75.12	68.53	68.85	68.65
7	9-6-1	89.79	94.82	91.41	92.29	93.01	92.69
8	9-6-1	70.99	75.92	72.55	78.53	79.82	78.89
9	9-6-1	65.30	66.85	65.98	68.63	69.87	68.52
10	9-6-1	52.99	55.30	53.86	57.18	57.82	55.48
11	9-6-1	49.34	46.24	47.89	42.12	48.53	45.18
12	9-9-1	39.34	42.06	41.47	38.18	38.82	37.47

Πίνακας 4.8. Πειραματικά αποτελέσματα της προτεινόμενης MLP αρχιτεκτονικής.

Τα αποτελέσματα που προέκυψαν από την προσομοίωση του MLP νευρωνικού δικτύου δείχνουν ότι το νευρωνικό δίκτυο με την καλύτερη επίδοση είναι το Νο. 7 με αρχιτεκτονική 10-6-1. Από την μελέτη όλων των πειραματικών αποτελεσμάτων του Πίνακα 4.8, προκύπτει ότι η συγκεκριμένη αρχιτεκτονική επιτυγχάνει ικανοποιητικά αποτελέσματα στην ταξινόμηση των εταιρειών με βάση τα δυο (2) επίπεδα αποδοτικότητας που ορίστηκαν στο εδάφιο 4.2. Η συμπεριφορά αυτή οφείλεται κυρίως στο μικρό αριθμό δειγμάτων που είναι διαθέσιμα για την εκπαίδευση των νευρωνικών δικτύων.

Τα αποτελέσματα του Πίνακα 4.8 αποδεικνύουν ότι όσο ο αριθμός των νευρώνων στο κρυφό επίπεδο αυξάνει, τόσο καλύτερη απόδοση επιτυγχάνει τα MLPs. Αυτή η συμπεριφορά είναι αναμενόμενη, δεδομένου ότι περισσότεροι νευρώνες στο κρυφό επίπεδο, συνεπάγονται μεγαλύτερο αριθμό βαρών και πολώσεων, με αποτέλεσμα να βελτιώνεται η ικανότητα αποθήκευσης επίκτητης γνώσης από το MLP. Η σύγκριση των επιδόσεων των MLPs με αρχιτεκτονικές 9-4-1 και 9-6-1 επιβεβαιώνει τον προαναφερθέντα ισχυρισμό. Ωστόσο, υπάρχει ένας μέγιστος αριθμός νευρώνων στο κρυφό επίπεδο, η υπέρβαση του οποίου έχει σαν αποτέλεσμα τη μείωση της απόδοσης του MLP.

Ο περιορισμός του αριθμού των κρυφών νευρώνων είναι αποτέλεσμα του φαινομένου overfitting, σύμφωνα με το οποίο το νευρωνικό δίκτυο έχει απομνημονεύσει τα δείγματα εκπαίδευσης, αλλά δεν έχει μάθει να γενικεύσει όταν στην είσοδό του παρουσιαστούν νέα δεδομένα [25]. Τα νευρωνικά δίκτυα από το 2^ο μέχρι και το 6^ο του Πίνακα 4.8 αντιμετωπίζουν μεγαλύτερη δυσκολία στην αναγνώριση των φυσιολογικών προτύπων και την ταξινόμησή τους στην κατάλληλη ομάδα, καθώς δεν έχουν τη δυνατότητα να μάθουν από τα δεδομένα εισόδου της φάσης εκπαίδευσης.

Ο Πίνακας 4.9 παρουσιάζει τα ληφθέντα αποτελέσματα των PNNs για την αποδοτικότητα των εταιρειών. Η radbas και η comper ήταν οι χρησιμοποιούμενες συναρτήσεις μεταφοράς για το κρυφό και για το επίπεδο εξόδου, αντίστοιχα. Ο αριθμός των νευρώνων για το επίπεδο εισόδου και το κρυφό επίπεδο των PNNs είναι σταθερός, ενώ ο αριθμός των νευρώνων εξόδου εξαρτάται από την κωδικοποίηση των επιθυμητών τιμών. Ο σκοπός του προτεινόμενου PNN είναι η κατάταξη των εισηγμένων στο Χρηματιστήριο Αξιών Αθηνών εταιρειών βάση της αποδοτικότητάς τους, με αποτέλεσμα η έξοδος του PNN να έχει 2 νευρώνες σύμφωνα με τα επίπεδα αποδοτικότητας που αναλύθηκαν στο εδάφιο 4.2. Συνεπώς, η μόνη παράμετρος που μπορεί να μεταβληθεί είναι η τιμή του παράγοντα εξάπλωσης (spread factor) της συνάρτησης ακτινικής βάσης του δεύτερου επιπέδου. Οι τιμές του παράγοντα εξάπλωσης παρουσιάζονται στη 2^η στήλη του Πίνακα 4.9. Τα λαμβανόμενα αποτελέσματα της 9-70-2 PNN τοπολογίας, μετά τη φάση προσομοίωσης φάσης, υπέστησαν παρόμοια επεξεργασία με εκείνα των MLPs. Ως εκ τούτου, η 3^η, 4^η και 5^η στήλη εμφανίζουν το ποσοστό επιτυχούς ταξινόμησης των προτύπων, ενώ η 6^η, 7^η και 8^η στήλη αποτυπώνουν την απόδοση των PNNs για τα δείγματα ανεπαρκούς αποδοτικότητας στα σύνολα ελέγχου και εκπαίδευσης και στο σύνολο δεδομένων, αντίστοιχα.

No.	spread	Ποσοστά Επιτυχούς Ταξινόμησης			Ποσοστά Επιτυχούς Ταξινόμησης		
		Σύνολο	Σύνολο	Σύνολο	Σύνολο	Σύνολο	Σύνολο
		Ελέγχου	Εκπαίδευσης	Δεδομένων	Ελέγχου	Εκπαίδευσης	Δεδομένων
1	0.2	29.90	97.12	69.00	32.84	87.82	71.09
2	0.5	30.80	98.97	73.01	32.84	92.70	73.42
3	0.8	30.80	98.97	73.01	32.84	92.70	73.42
4	2.7	36.72	62.12	53.32	41.98	65.20	56.21
5	3.6	39.48	58.62	49.71	44.82	58.63	52.37
6	8.4	93.20	95.24	95.87	91.10	94.21	92.72
7	17.7	43.15	45.82	44.03	44.82	47.97	47.02
8	25.4	38.50	39.93	38.95	43.11	44.94	44.05
9	40.0	29.90	31.28	30.05	34.81	36.87	35.81
10	50.0	27.98	29.12	28.10	20.18	25.16	22.91

Πίνακας 4.9. Πειραματικά αποτελέσματα της προτεινόμενης PNN αρχιτεκτονικής.

Μια μικρή τροποποίηση του παράγοντα εξάπλωσης της ακτινικής συνάρτησης βάσης δεν επηρεάζει την απόδοση των PNNs, όπως φαίνεται από τα Νο. 1, Νο. 2 και Νο. 3 νευρωνικά δίκτυα του Πίνακα 4.9 Η διαφορά μεταξύ 0.2, 0.5 και 0.8 είναι ανεπαίσθητη, οπότε δεν υπάρχει διαφοροποίηση επί των ληφθέντων αποτελεσμάτων. Επισημαίνεται ότι το ποσοστό της επιτυχούς ταξινόμησης για τα σύνολα δεδομένων και ελέγχου αυξάνει καθώς αυξάνεται η τιμή του συντελεστή εξάπλωσης. Ωστόσο, υπάρχει μια τιμή του συντελεστή εξάπλωσης που αποτελεί όριο στη βελτίωση της απόδοσης για το σύνολο ελέγχου. Το όριο για την παράμετρο εξάπλωση των υλοποιημένων PNNs ισούται με 8.4, ενώ τιμές μεγαλύτερες από τον παραπάνω αριθμό μειώνουν την ικανότητα πρόβλεψης των PNNs για τον σύνολο ελέγχου. Τα Νο. 1, Νο. 2 και Νο. 3 PNNs

δεν έχουν επαρκή ικανότητα γενίκευσης, σε αντίθεση με το Νο. 6 PNN, το οποίο ταξινομεί σωστά νέα άγνωστα πρότυπα που παρουσιάζονται στο επίπεδο εισόδου του νευρικού δικτύου.

Επισημαίνεται ότι Νο. 5 PNN υπερτερεί του Νο. 4 PNN, όπως συμπεραίνεται από τα αποτελέσματα που λαμβάνονται. Παρόλα αυτά, η σύγκριση των διαφορών μεταξύ της επιτυχούς ταξινόμησης των προτύπων για τα δεδομένα ελέγχου, εκπαίδευσης και το σύνολο δεδομένων είναι εξαιρετικά σημαντική. Το Νο. 5 PNN διακρίνει τα δείγματα εκπαίδευσης ικανοποιητικά και αυτό είναι ένα χαρακτηριστικό που προσθέτει στην ικανότητα ταξινόμησης, ενώ την ίδια στιγμή η απόδοσή του μειώνεται ελαφρώς, εάν συγκριθεί με το Νο. 4 PNN. Το Νο. 5 PNN παρουσιάζει συμπεριφορά overfitting. Κατά συνέπεια, το Νο. 4 PNN υπερτερεί του Νο. 5 PNN, όσον αφορά την ικανότητα γενίκευσης. Συνεπώς το νευρωνικό δίκτυο που είναι κατάλληλο για την ταξινόμηση των εταιρειών από πλευράς αποτελεσματικότητας είναι το PNN με spread ίσο με 8.4.

ΒΙΒΛΙΟΓΡΑΦΙΑ ΚΕΦΑΛΑΙΟΥ

1. Altman E., “Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy”, *Journal of Finance*, 1968, Vol 23, pp. 589-609.
2. Atiya A., “Bankruptcy Prediction for Credit Risk Using Neural Networks: A Survey and New Results”, *IEEE Transactions on Neural Networks*, 2001, Vol. 12, pp. 929-935.
3. Beaver W.H., “Financial Ratios as Predictors of Failure, Empirical Research in Accounting: Selected Studies”, *Journal of Accounting Research*, 1966, Vol. 4, pp. 71-111.
4. Dokas I., “Can Macroeconomic Variables Enhance the Prediction Power of Accounting-Based Financial Failure Prediction Models? Evidence from the Greek Economy”, *International Journal of Economics*, 2009, Vol. 3, No. 1, pp 1-12.
5. Messier W.F, and Hansen J.V., “Inducing Rules for Expert System Development: An Example Using Default and Bankruptcy Data”, *Management Science*, 1988, Vol. 34, No 12, 1403-1415.
6. Saleem Q., and Rehman U., “Impacts of Liquidity Ratios on Profitability (Case of Oil and Gas Companies of Pakistan)”, *Interdisciplinary Journal of Research in Business*, Vol. 1, No. 7, pp. 95-98.
7. Papik K., Molnar B., Schaefer R., Dombovari Z., Tulassay Z., and Feher J. “Application of Neural Networks in Medicine - A Review”, *Med Sci Monit.*, 1998, pp.538-546.
8. Bilbao J., and Bravo E., “Comparing Newton with ANN for Load Flow”, *WSEAS Transactionson Mathematics*, 2004, Vol. 3, pp. 31-36.
9. Tafeit E., and Reibnegger G., “Review Artificial Neural Networks in Laboratory Medicine and Medical Outcome Prediction”, *Clin Chem Lab Med.*, 199, Vol. 37, pp. 845-853.
10. Shieh J., Fan S., and Shi W., “The Intelligent Model of a Patient Using artificial Neural Networks for Inhalational Anaesthesia”, *Journal of the Chinese Institute of Chemical Engineers*, 2002, Vol. 33, No. 6, pp. 609-620.
11. Piuri V., and Scotti F., “Morphology Classification of Blood Leucocytes by Microscope Images”, *IEEE International Conference on Computational Intelligence International Conference on Image, Speech and Signal Analysis*, 1992, pp. 530-533.
12. Orr R., “Use of a Probabilistic Neural Network to Estimate the Risk of Mortality After Cardiac Surgery”, 1997, *J. Medical Decision Making*, Vol. 17, pp. 178-185.
13. Parzen E., “On Estimation of a Probability Density Function and Mode”, *Annals of Mathematical Statistics*, 1962, Vol. 33, No. 3, pp. 1065-1076.

14. Chung, T., Chen Y.M., and Tang S.C., “A Hybrid System Integrating Signal Analysis and Probabilistic Neural Network for User Motion Detection in Wireless Networks”, *Expert Systems with Applications*, 2010, Vol. 39, pp. 3392-3403.
15. Iliadis L., “Intelligent Information Systems and Applications in Risk Estimation”, *Stamoulis Publishing*, Greece, 2007.
16. Xu C., Zhang H., Huang C., and Peng D., “Study of Fault Diagnosis Based on Probabilistic Neural Network for Turbine Generator Unit”, *International Conference on Artificial Intelligence and Computational Intelligence (AICI)*, 2010, Vol. 1, pp. 275-279.
17. Burrascano P., “Learning Vector Quantization for the Probabilistic Neural Network”, *IEEE Trans Neural Net.*, 1991, pp. 458-461.
18. Specht D., “Probabilistic Neural Networks”, *Neural Networks*, 1990, Vol. 3, pp. 109-118.
19. Rao P., Devi T., Kaladhar D., Sridhar G., and Rao A., “A Probabilistic Neural Network Approach for Protein Superfamily Classification”, *Journal of Theoretical and Applied Information Technology*, 2009, pp. 101-105.
20. Mantzaris D., Anastassopoulos G., Adamopoulos A., and Gardikis S., “A Non- Symbolic Implementation of Abdominal Pain Estimation in Childhood”, *Information Science*, 2008, Vol. 178, pp. 3860-3866.
21. Othman M.F., and Basri M.A.M., “Probabilistic Neural Network for Brain Tumor Classification”, *Second Int. Conference on Intelligent Systems, Modelling and Simulation (ISMS)*, 2011, pp. 136-138.
22. Anderson D., and McNeil G., “Artificial Neural Networks Technology”, *Kaman Science Corporation*, 1992.
23. Beale M., Hagan M., and Demuth H., “Neural Network Toolbox User’s Guide”, *The Math Works Inc*, 2013.
24. Mantzaris D., Anastassopoulos G., and Adamopoulos A., “Genetic Algorithm Pruning of Probabilistic Neural Networks in Medical Disease Estimation”, *Neural Networks*, 2011, Vol. 24, pp. 831-835.
25. Mantzari V., and Mantzaris D., “Solar Radiation: Cloudiness Forecasting Using a Soft Computing Approach”, *Artificial Intelligence Research*, 2013, Vol. 2, pp. 69-80.

ΣΥΜΠΕΡΑΣΜΑΤΑ ΠΡΟΟΠΤΙΚΕΣ

5

Η σύσταση εταιρειών σε μια χώρα αποτελεί θετική ένδειξη για την οικονομία της καθώς συνεπάγεται καινούριες επενδύσεις, νέες θέσεις εργασίας και προσέλκυση, ενδεχομένως, ξένων κεφαλαίων. Η δραστηριοποίηση πολλών εταιρειών σε μια χώρα, την καθιστά ανεπτυγμένη και η οικονομία της λειτουργεί με βάση την ελεύθερη αγορά και το θεμιτό ανταγωνισμό. Πέρα από τη δημιουργία της επιχείρησης, ιδιαίτερη έμφαση χρειάζεται να δοθεί στη λειτουργία της. Απαιτείται συνεχής έλεγχος τόσο της παραγωγικής διαδικασίας όσο και της οικονομική της πορείας.

Η χρηματοοικονομική ανάλυση δίνει τη δυνατότητα στην επιχείρηση να εντοπίσει αδυναμίες της από οικονομικής πλευράς ενώ παράλληλα αποτελεί εργαλείο παρακολούθησης της πορείας της. Είναι σαφές, ότι η κατάσταση της επιχείρησης αποτυπώνεται στις χρηματοοικονομικές της καταστάσεις. Μια ανάλυση που βασίζεται στις συγκεκριμένες καταστάσεις είναι απαραίτητη για την αξιολόγηση της αποτελεσματικότητας της επιχείρησης. Η έγκαιρη και έγκυρη χρηματοοικονομική παρακολούθηση της επιχείρησης αποτελεί προτεραιότητα τόσο της διοίκησής όσο και των μετόχων της. Αυτό έχει σαν αποτέλεσμα την επιτακτική ανάπτυξη συστημάτων που βοηθούν τους ειδικούς στην αποτελεσματικότερη αξιολόγηση της επιχείρησης.

Στην παρούσα διατριβή αναπτύχθηκαν ευφυείς υπολογιστικές μέθοδοι και αξιολογήθηκε η συνεισφορά τους στην παρακολούθηση της αποτελεσματικότητας των επιχειρήσεων. Με οδηγό την αξιοποίηση των τεχνολογιών της πληροφορικής και των επικοινωνιών στη χρηματοοικονομική, έγινε εφαρμογή όλων των σταδίων της μηχανικής μάθησης, από την αρχική επεξεργασία των δεδομένων μέχρι την ανάπτυξη ενός πρωτότυπου συστήματος για την αυτοματοποίηση της αξιολόγησης της αποτελεσματικότητας της επιχείρησης.

Η υλοποίηση των ευφών υπολογιστικών μεθόδων αλγορίθμων βασίστηκε σε μοντέλα τεχνητών νευρωνικών δικτύων. Συγκεκριμένα, αναπτύχθηκαν νευρωνικά δίκτυα για την αποτίμηση της κατάστασης μιας επιχείρησης. Χαρακτηριστικό τους γνώρισμα είναι μπορούν να επιτύχουν εκτίμηση της χρηματοοικονομικής κατάστασης μιας επιχείρησης χωρίς να υπάρχουν περιορισμοί, χρησιμοποιώντας τόσο αριθμοδείκτες όσο και ποσά από του ισολογισμού.

Τα αποτελέσματα που προέκυψαν από τα προτεινόμενα μοντέλα, στα πλαίσια της παρούσας διατριβής, αποδεικνύουν την ικανότητα εφαρμογής των νευρωνικών δικτύων σε χρηματοοικονομικούς τομείς που δεν έχουν χρησιμοποιηθεί μέχρι σήμερα. Αυτό τα καθιστά ικανά ώστε να μπορούν να εφαρμοστούν και σε άλλους τομείς της οικονομίας, αποτελώντας αξιόπιστες ευφυείς τεχνικές εκτίμησης.

Μια πρώτη σημαντική εργασία που μπορεί να λάβει χώρα στο πλαίσιο επέκτασης της παρούσας ερευνητικής προσπάθειας είναι η μελέτη, υλοποίηση και ενσωμάτωση μιας αυτοματοποιημένης διαδικασίας εκπαίδευσης των ευφών υπολογιστικών μεθόδων στα νέα διαθέσιμα δεδομένα με σκοπό τη βελτιστοποίηση της απόδοσής τους και της ικανότητας γενίκευσης στα καινούρια οικονομικά στοιχεία.

ΓΛΩΣΣΑΡΙ

ANN	Artificial Neural Network
BFGS	Broyden, Fletcher, Goldfarb και Shanno
CI	Computational Intelligence
GRNN	Generalized Regression Neural Network
LVQ	Learning Vector Quantization
MLP	Mult-layer Perceptron
MSE	Mean Square Error
OSS	One Step Secant
PDF	Probabilistic Density Function
PNN	Probabilistic Neural Network
RBF	Radial Basis Function
SOFM	Self-Organizing Feature Map
VLSI	Very Large Scale Integration
TN	Τεχνητή Νοημοσύνη
TNΔ	Τεχνητά Νευρωνικά Δίκτυα
ΥΝ	Υπολογιστική Νοημοσύνη