



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ
ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

Διαχείριση Δομής Νευροασαφών Συστημάτων με Ανάλυση Ευαισθησίας

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ναταλί- Μαρία Θ. Ραουζαίου

Επιβλέπων : Ανδρέας- Γεώργιος Ν. Σταφυλοπάτης
Καθηγητής Ε.Μ.Π

Αθήνα, Οκτώβριος 2008



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ
ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

Διαχείριση Δομής Νευροασαφών Συστημάτων με Ανάλυση Ευαισθησίας

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ναταλί- Μαρία Θ. Ραουζαίου

Επιβλέπων : Ανδρέας- Γεώργιος Ν. Σταφυλοπάτης
Καθηγητής Ε.Μ.Π

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 20^η Οκτωβρίου 2008.

.....

Ανδρέας- Γεώργιος
Σταφυλοπάτης
Καθηγητής ΕΜΠ

.....

Στέφανος Κόλλιας
Καθηγητής ΕΜΠ

.....

Γεώργιος Στάμου
Λέκτορας ΕΜΠ

Αθήνα, Οκτώβριος 2008

Ναταλί- Μαρία Θ. Ραουζαίου

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Ναταλί- Μαρία Θ. Ραουζαίου

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς το συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περιεχόμενα

Περιεχόμενα.....	5
Σχήματα.....	7
Πίνακες.....	9
Πρόλογος.....	11
Περίληψη	13
Abstract	15
Κεφάλαιο 1: Εισαγωγή.....	17
1.1 Γενικά	17
1.2 Περιβαλλοντική Ευφυΐα- Νευροασαφή Συστήματα.....	18
1.3 Περιγραφή Διπλωματικής.....	21
1.3.1 Σκοπός – Κίνητρα	21
1.3.2 Διάρθρωση Διπλωματικής.....	23
Κεφάλαιο 2:Νευροασαφή συστήματα γινομένου – Ανάλυση Ευαισθησίας.....	25
2.1 Γενικά	25
2.2 Μοντελοποίηση νευροασαφών συστημάτων και προσδιορισμός δομής.....	26
2.3 Κίνητρο και Συμβολή.....	29
Κεφάλαιο 3: Μεθοδολογία- Περιγραφή Τεχνικής.....	31
3.1 Γενικά	31
3.2 Supfunis	32
3.3 Αρχιτεκτονική Τύπου RAN (Resource Allocating Network)	33
3.4 Aranfis- Μοντέλο Προσαρμοζόμενης Δομής με Ανάθεση Πόρων.....	36
3.4.1 Διαδικασία Μάθησης για το ARANFIS.....	37
3.4.2 Προσαρμογή ρυθμού μάθησης:	39

| Περιεχόμενα

3.5	Αξιολόγηση και Δυναμική επιλογή κανόνων- Προτεινόμενη Μέθοδος	40
3.6	Αξιολόγηση και Δυναμική Επιλογή Κανόνων στο μοντέλο ARANFIS.....	41
3.7	Περιγραφή Λειτουργίας του μοντέλου ARANFIS με την προτεινόμενη μεθοδολογία	45
	Κεφάλαιο 4: Πειραματική Μελέτη.....	47
	Κεφάλαιο 5: Συμπεράσματα– Μελλοντικές επεκτάσεις.....	63
5.1	Συμπεράσματα.....	63
5.2	Μελλοντικές Επεκτάσεις.....	65
	Παράρτημα Α:	67
	Παράρτημα Β:.....	71
	Βιβλιογραφία	73
	Βιογραφικό Σημείωμα	77

Σχήματα

Σχήμα 3.1: Παράδειγμα δικτύου με αρχιτεκτονική τύπου RAN	34
Σχήμα 3.2: Αρχιτεκτονική ARANFIS	37
Σχήμα 3.3: Ιεραρχική δομή κανόνων- Χωρισμός σε δυο υποσύνολα	44
Σχήμα 4.1: Απόδοση εκπαιδευμένου συστήματος για το σύνολο Breast Cancer	56
Σχήμα 4.2: Πλήθος κανόνων με ισχύ ενεργοποίησης > 1 για το σύνολο Breast Cancer	57
Σχήμα 4.3: Απόδοση συστήματος μόνο με την κύρια βάση κανόνων για το σύνολο Breast Cancer	57
Σχήμα 4.4: Απόδοση εκπαιδευμένου συστήματος για το σύνολο Ionosphere	58
Σχήμα 4.5: Πλήθος κανόνων με ισχύ ενεργοποίησης > 1 για το σύνολο Ionosphere	58
Σχήμα 4.6: Απόδοση συστήματος μόνο με την κύρια βάση κανόνων για το σύνολο Ionosphere	59
Σχήμα 4.7: Απόδοση εκπαιδευμένου συστήματος για το σύνολο Pima Indians	59
Σχήμα 4.8: Πλήθος κανόνων με ισχύ ενεργοποίησης > 1 για το σύνολο Pima Indians	60
Σχήμα 4.9: Απόδοση συστήματος μόνο με την κύρια βάση κανόνων για το σύνολο Pima Indians	60
Σχήμα 4.10: Απόδοση εκπαιδευμένου συστήματος για το σύνολο Vehicle	61
Σχήμα 4.11: Πλήθος κανόνων με ισχύ ενεργοποίησης > 1 για το σύνολο Vehicle	61

Σχήμα 4.12: Απόδοση συστήματος μόνο με την κύρια βάση κανόνων για το σύνολο Vehicle

..... 62

Πίνακες

Πίνακας 4.1: Αποτελέσματα πειραμάτων για την εκπαίδευση του συστήματος	49
Πίνακας 4.2: Αποτελέσματα απλής προσομοίωσης εκπαιδευμένου συστήματος	50
Πίνακας 4.3: Αποτελέσματα προσομοίωσης εκπαιδευμένου συστήματος με ιεράρχηση κανόνων με κριτήριο την ευαισθησία.	50
Πίνακας 4.4: Συγκριτικός πίνακας για τις δυο προσομοιώσεις: Το σύστημα χωρίς δυναμική αξιολόγηση και ιεράρχηση κανόνων συγκρίνεται με το σύστημα όπου εφαρμόστηκε η προτεινόμενη μέθοδος δυναμικής αξιολόγησης κανόνων και χωρισμού τους σε δυο υποσύνολα.	51
Πίνακας 4.5: Αποτελέσματα πειραμάτων για την εκπαίδευση του συστήματος με κριτήρια προσθήκης κανόνων ϵ_2 και δ	52
Πίνακας 4.6: Αποτελέσματα απλής προσομοίωσης εκπαιδευμένου συστήματος με κριτήρια ϵ_2 και δ	52
Πίνακας 4.7: Αποτελέσματα προσομοίωσης εκπαιδευμένου συστήματος με ιεράρχηση κανόνων με κριτήριο την ευαισθησία και κριτήρια προσθήκης κανόνων δ και ϵ_2	53
Πίνακας 4.8: Συγκριτικός πίνακας για τις δυο προσομοιώσεις.....	53
Πίνακας 4.9: Αποτελέσματα πειραμάτων για την εκπαίδευση του συστήματος με κριτήριο προσθήκης	54
Πίνακας 4.10: Αποτελέσματα απλής προσομοίωσης εκπαιδευμένου συστήματος με κριτήριο	54

| Πίνακες

Πίνακας 4.11: Αποτελέσματα προσομοίωσης εκπαιδευμένου συστήματος με ιεράρχηση κανόνων με κριτήριο την ευαισθησία και κριτήριο προσθήκης κανόνων ϵ_2 54

Πίνακας 4.12: Συγκριτικός πίνακας για τις δυο προσομοιώσεις 55

Πρόλογος

Ο χώρος των νευρωνικών δικτύων αρχικά, και αργότερα των νευροασαφών συστημάτων, μου κέντριζε πάντα το ενδιαφέρον, όχι μόνο λόγω της πληθώρας των εφαρμογών στις οποίες χρησιμοποιείται, αλλά κυρίως γιατί πρόκειται για ένα χώρο που διαρκώς εξελίσσεται και πρωταγωνιστεί στη σύγχρονη τεχνολογική πραγματικότητα. Αυτός είναι και ο λόγος που θέλησα να ασχοληθώ με τα νευροασαφή συστήματα και να αναλάβω τη συγκεκριμένη διπλωματική εργασία.

Η εκπόνηση της εργασίας πραγματοποιήθηκε υπό την επίβλεψη του καθηγητή κ. Ανδρέα-Γεώργιου Σταφυλοπάτη, τον οποίο ευχαριστώ θερμά για την εμπιστοσύνη που έδειξε στο πρόσωπό μου, την υπομονή και τη διάθεση συνεργασίας σε όλη τη διάρκεια της έρευνάς μου. Οι γνώσεις και η εμπειρία του, σε συνδυασμό με το ενδιαφέρον και την προθυμία του να μου προσφέρει βοήθεια όπου τη χρειάστηκα, συνέβαλαν αποφασιστικά στην ανάπτυξη και ολοκλήρωση της παρούσας διπλωματικής.

Ιδιαίτερες ευχαριστίες οφείλω και στο Μηνά Περτσσελάκη, διδάκτορα ΕΜΠ, η διδακτορική διατριβή του οποίου αποτέλεσε έναυσμα για τη δική μου διπλωματική εργασία. Η ολοκληρωμένη και εμπειριστατωμένη δουλειά του, στην οποία στηρίχτηκε η έρευνά μου, και το ενδιαφέρον του για νέες ιδέες και προτάσεις, με έκαναν να αγαπήσω ακόμα περισσότερο το αντικείμενο. Η συνεργασία μας υπήρξε άψογη καθ' όλη τη διάρκεια της εκπόνησης της εργασίας, και η βοήθειά του υπήρξε πολύτιμη τόσο σε επιστημονικό όσο και σε προσωπικό επίπεδο.

Θέλω, τέλος, να ευχαριστήσω και την οικογένεια μου, αλλά και το φιλικό μου περιβάλλον για τη στήριξη και τη βοήθειά τους σε όλη τη διάρκεια αυτής της έρευνας, μέχρι και την ολοκλήρωση της εργασίας.

Ραουζαίου Ναταλί- Μαρία
Αθήνα, Οκτώβριος 2008

Περίληψη

Οι τομείς των νευροασαφών συστημάτων και της περιβαλλοντικής ευφυΐας είναι από τους πιο ενδιαφέροντες και ταχύτερα αναπτυσσόμενους τομείς της σύγχρονης επιστημονικής πραγματικότητας, που βρίσκουν εφαρμογή σε μια πληθώρα καθημερινών δραστηριοτήτων. Οι ανάγκες που πρέπει να καλυφθούν διαρκώς αυξάνονται, δημιουργώντας έτσι νέες απαιτήσεις που καλούνται να ικανοποιήσουν τα νευροασαφή συστήματα. Στην παρούσα διπλωματική εργασία γίνεται προσπάθεια να αντιμετωπιστούν κάποιες από τις αδυναμίες που εμφανίζουν τα συστήματα αυτά, και ταυτόχρονα να βελτιωθούν τα χαρακτηριστικά τους, ώστε να διευρυνθεί ακόμα περισσότερο το φάσμα των εφαρμογών τους. Στόχος είναι να μπορέσουν να καλύψουν ανάγκες όπως η εξοικονόμηση ενέργειας και η μείωση του σχεδιαστικού, υπολογιστικού και λειτουργικού τους κόστους, αλλά ταυτόχρονα να αυξηθεί, ή έστω να μη μειωθεί, η απόδοσή τους. Η μέθοδος που εφαρμόζεται κάνει χρήση ενός συγκεκριμένου μοντέλου νευροασαφών συστημάτων, των νευροασαφών συστημάτων γινομένου, τα οποία συνδυάζονται με ένα πολύ γνωστό στατιστικό και ερευνητικό εργαλείο, την ανάλυση ευαισθησίας. Χρησιμοποιείται ένα μοντέλο προσαρμοζόμενης δομής με ανάθεση πόρων (ARANFIS), στο οποίο εφαρμόζεται η ανάλυση ευαισθησίας στο επίπεδο των κρυφών κόμβων κανόνων και όχι στο επίπεδο της εισόδου, όπως γίνεται μέχρι σήμερα. Αυτή είναι και η καινοτομία που προσπαθεί να εισάγει η παρούσα διπλωματική εργασία στο χώρο των νευροασαφών συστημάτων: δυναμική διαχείριση της δομής και των πόρων ενός δικτύου, παρέμβαση στην εσωτερική του λειτουργία και δομή, με στόχο το χαμηλότερο συνολικό κόστος αλλά την ίδια (ή και καλύτερη) απόδοση. Η μέθοδος που προτείνεται βασίζεται στην αξιολόγηση και ιεράρχηση του τελικού αριθμού των κανόνων, που έχει τη δυνατότητα να προσθέσει το δίκτυο κατά την εκπαίδευση, σε δυο υποσύνολα. Η αξιολόγηση αυτή γίνεται με βάση το πόσο σημαντικός είναι κάθε κανόνας για την έξοδο του δικτύου. Υπολογίζεται, δηλαδή, με την ανάλυση ευαισθησίας, πόση επίδραση έχει κάθε κρυφός κόμβος κανόνας στην έξοδο του συστήματος, και ανάλογα κατατάσσεται στο κύριο ή το δευτερεύον υποσύνολο κανόνων. Το σύστημα χρησιμοποιεί το κύριο υποσύνολο κανόνων ή και τα δυο, ανάλογα με ένα κατώφλι εμπιστοσύνης. Η διαδικασία αυτή διατηρεί την ίδια απόδοση στο σύστημα, αλλά το καθιστά οικονομικότερο. Αυτό επιβεβαιώνεται και από μια σειρά δοκιμών με τέσσερα διαφορετικά σύνολα δεδομένων, των οποίων τα αποτελέσματα παρατίθενται στην παρούσα διπλωματική εργασία.

Λέξεις-Κλειδιά:

Νευροασαφή συστήματα γινομένου, Ανάλυση ευαισθησίας, Δυναμική διαχείριση δομής, Προσαρμοζόμενη δομή, Δυναμική ιεράρχηση κανόνων, Κύριο και δευτερεύον υποσύνολο κανόνων.

Abstract

Neurofuzzy systems and ambient intelligence are two of the most interesting and rapidly developed fields of actual technology, applied in many activities of everyday human life. As time passes by, new needs arise and, as a result, neurofuzzy systems have to meet people's and scientists' fresh requirements. In this diploma thesis, we try to correct some of the drawbacks of neurofuzzy systems and ameliorate their features, in order to broaden their spectrum of applications. Fresh needs, such as energy saving and the reduction of designing, running and computational cost of neurofuzzy systems, have to be satisfied. Meanwhile, their performance has to maintain the same, or even better, has to be improved. In this diploma thesis, we use a certain model of neurofuzzy systems, neurofuzzy product systems, combined with a very popular and commonly used statistical and mathematical tool, sensitivity analysis. ARANFIS, a resource allocating neurofuzzy inference system, is used, and sensitivity analysis is applied between the level of hidden nodes (rules) and the system output. This is different from the usual approach, according to which sensitivity analysis is applied between the inputs and the outputs of the network. In the field of neurofuzzy networks, this approach constitutes a novelty: we try to change the internal structure of the network, manage its resources dynamically, in order to build an ideal network with low total cost and high performance. The method suggested is based on the evaluation of rules that the network adds during its training. The rules are divided in two subsets, the main subset and the secondary one, depending on their importance for the network output. Using sensitivity analysis, we can estimate the effect of each rule on the system output and, therefore, define which rule belongs to the main subset or not. The system can use either the main subset, or both of them, so as to satisfy a certain confidence measure. As a result, we have a system of the same performance, at a lower total cost. This can be confirmed by the results of tests included in this diploma thesis, performed using a variety of datasets.

Keywords:

Neurofuzzy product systems, Sensitivity analysis. Dynamic structure management, Resource allocating, Dynamic rules hierarchy, Main and secondary rules subset.

Κεφάλαιο 1: Εισαγωγή

1.1 Γενικά

Ο άνθρωπος, έχοντας ξεπεράσει εδώ και δεκαετίες τους ενδοιασμούς και τους φόβους του απέναντι στα εκατοντάδες σύγχρονα επιτεύγματα της τεχνολογίας, μαθαίνει να ζει με την τεχνολογία και να εκμεταλλεύεται αυτά που του προσφέρει για να κάνει την καθημερινότητά του πιο εύκολη και πιο ευχάριστη. Προς αυτήν την κατεύθυνση στρέφονται και οι προσπάθειες των επιστημονικών κοινοτήτων: να δημιουργήσουν ένα νέο τρόπο αλληλεπίδρασης του ανθρώπου με το περιβάλλον και με την τεχνολογική πραγματικότητα, ο οποίος θα του καλύπτει με μεγαλύτερη επιτυχία τις ανάγκες του, εξαλείφοντας μάλιστα όλο και περισσότερο τις δυσκολίες και τα προβλήματα που προκύπτουν από την αλληλεπίδραση του ανθρώπου με «ψυχρά» μηχανήματα. Είναι λοιπόν κοινώς αποδεκτό ότι το στοιχείο που χρειάζεται να προστεθεί στις «άψυχες» μηχανές είναι η προσαρμοστικότητα, η ικανότητα να μπορούν να αντιδρούν στα διάφορα ερεθίσματα δυναμικά, η δυνατότητα δηλαδή να «σκέφτονται» και να λαμβάνουν αποφάσεις ανάλογα με τις εκάστοτε συνθήκες και απαιτήσεις του χρήστη, και όχι στατικά και μονοσήμαντα.

1.2 Περιβαλλοντική Ευφυΐα- Νευροασαφή Συστήματα

Περιβαλλοντική ευφυΐα

Προς αυτήν την κατεύθυνση κινείται και η έννοια της περιβαλλοντικής ευφυΐας. Στόχος είναι η δημιουργία ενός τεχνολογικού περιβάλλοντος με δείκτη ευφυΐας, ενός ψηφιακού περιβάλλοντος πληροφoρίας, το οποίο ευαισθητοποιείται, προσαρμόζεται και ανταποκρίνεται στην παρουσία ανθρώπων, και έχει ως στόχο να δημιουργήσει την επιθυμητή ατμόσφαιρα και λειτουργικότητα με τη χρήση ευφών συστημάτων και την προσφορά προσωποποιημένων υπηρεσιών. Η περιβαλλοντική ευφυΐα χαρακτηρίζεται από τα εξής βασικά στοιχεία: πανταχού παρουσία, διαφάνεια και νοημοσύνη.

Η πανταχού παρουσία (ubiquity) περιγράφει την κατάσταση κατά την οποία παντού γύρω μας και σε οποιαδήποτε χρονική στιγμή, υπάρχει ένα πλήθος διασυνδεδεμένων συσκευών. Η διαφάνεια (transparency) προϋποθέτει ότι οι συσκευές αυτές είναι άορατες στο χρήστη και λειτουργούν ανεξάρτητα σχεδόν στον περιβάλλοντα χώρο. Τέλος η νοημοσύνη (intelligence) είναι η μορφή ανώτερης ευφυΐας που απαιτείται να διαθέτουν αυτές οι συσκευές. Συχνά χρησιμοποιείται και ο όρος αντίληψη περιεχομένου (context-awareness) [1] αντί για τον όρο νοημοσύνη, αφού μιλάμε για την ικανότητα μιας ευφούς συσκευής να αντιλαμβάνεται το χώρο και τα άτομα που την περιβάλλουν (χρήστες), και να προσαρμόζει σε αυτά τις «αντιδράσεις» και τη λειτουργία της. Ως περιεχόμενο μπορεί να θεωρηθεί ακόμα και η εσωτερική κατάσταση της ίδιας της συσκευής, που μπορεί να αποτελεί άλλο ένα κριτήριο για τον τρόπο λειτουργίας της.

Θα μπορούσαμε απλά να πούμε ότι η προσπάθεια που γίνεται είναι να κατασκευαστούν συσκευές που θα μπορούν να αντιμετωπίζουν τα προβλήματα με τρόπο ανθρώπινο, θα εμφανίζουν δηλαδή ανθρώπινα χαρακτηριστικά, με πιο βασικό αυτό της προσαρμοστικότητας στις εκάστοτε συνθήκες και περιπτώσεις.

Πληθώρα παραδειγμάτων μπορούμε να βρούμε στις καθημερινές οικιακές ηλεκτρικές και ηλεκτρονικές συσκευές. «Εξυπνα» πλυντήρια που μπορούν να αποφασίσουν για τον τρόπο πλύσης αναγνωρίζοντας τα χαρακτηριστικά των ρούχων, κινητά τηλέφωνα που θα προσαρμόζουν τη λειτουργία τους στο περιβάλλον στο οποίο βρίσκεται ο χρήστης χωρίς κάποια δική του παρέμβαση και θα του παρέχουν υπηρεσίες εξατομικευμένες, που θα απευθύνονται στις δικές του ανάγκες, ανά πάσα στιγμή, ευφυή αυτοκίνητα που θα αποτρέπουν τα ατυχήματα και θα σώζουν ανθρώπινες ζωές, και πολλά άλλα παραδείγματα που ήδη έχουν κάνει την εμφάνισή τους στην καθημερινότητά μας.

Κοινό χαρακτηριστικό για όλα αυτά τα συστήματα, η «ευφυΐα». Ευφυή συστήματα είναι τα συστήματα που επιδεικνύουν λογική, εμπειρική μάθηση, και ικανότητες λήψης λογικών αποφάσεων χωρίς τη μεσολάβηση του ανθρώπου, σχεδιασμένα όμως έχοντας ως πρότυπο τον τρόπο που οι άνθρωποι και οι άλλοι βιολογικοί οργανισμοί εξελίσσονται, και αναπτύσσουν ικανότητες επίλυσης δύσκολων προβλημάτων. Στο σχεδιασμό τους χρησιμοποιούνται παραδοσιακές τεχνικές επεξεργασίας αριθμητικών- συμβολικών

αναπαραστάσεων της γνώσης, εξελικτικοί αλγόριθμοι, ασαφή και νευροασαφή συστήματα, τα οποία παρέχουν μια μεθοδολογία χειρισμού ασαφούς- μη ακριβούς πληροφορίας, προσφέροντας έτσι πολλές δυνατότητες εφαρμογής τους.

Από τα μέσα της δεκαετίας του '90 μέχρι σήμερα, η βασική πρόκληση στο πεδίο των ευφών συστημάτων, έχει να κάνει με τη δημιουργία αυτόνομων ευφών οντοτήτων, γνωστών και ως ευφείς πράκτορες (intelligent agents), που θέλουν να μιμηθούν την ανθρώπινη συμπεριφορά στην αντιμετώπιση καθημερινών καταστάσεων, επίλυση προβλημάτων και απόκτηση γνώσης, ώστε να διευκολυνθεί κάθε ανθρώπινη δραστηριότητα [13].

Οι σύγχρονες όμως επιστημονικές μελέτες επεκτείνονται και πέρα από αυτό. Έχοντας ήδη καταφέρει να δημιουργήσουμε ευφυή συστήματα, το ζητούμενο πλέον είναι να τα κάνουμε πιο «ευαίσθητα» σε ερεθίσματα, και πιο ευπροσάρμοστα σε καταστάσεις, ώστε η συμπεριφορά και οι αντιδράσεις τους να είναι όσο το δυνατόν πιο ανθρώπινες. Προς αυτήν την κατεύθυνση προσανατολίζονται και οι νεότερες έρευνες, που θέλουν τα ευφυή συστήματα όχι μόνο να δέχονται πληροφορίες και να εξάγουν συμπεράσματα με βάση κάποιους δεδομένους κανόνες, αλλά και να μπορούν να κρίνουν και να αξιολογούν αυτά τα δεδομένα και τους κανόνες, και να αποφασίζουν ποια πληροφορία θα χρησιμοποιήσουν και πώς.

Νευροασαφή Συστήματα

Ένα από τα πιο σύνθετα και «μυστήρια» συστήματα του ανθρώπινου οργανισμού είναι το νευρικό σύστημα, υπεύθυνο για τη συμπεριφορά και για τις αντιδράσεις κάθε βιολογικού οργανισμού στα ερεθίσματα του περιβάλλοντός του. Αυτά ακριβώς τα βιολογικά νευρωνικά δίκτυα και την υπολογιστική τους ικανότητα, επιδιώκουν να μοντελοποιήσουν τα τεχνητά νευρωνικά δίκτυα (ΤΝΔ), με στόχο τη μάθηση και την αποτύπωση της γνώσης, όπως συμβαίνει και με τους βιολογικούς οργανισμούς. Τα τεχνητά νευρωνικά δίκτυα, για να μπορέσουν να αποτυπώσουν την επιθυμητή γνώση, έχουν ανάγκη από εκπαίδευση, από μια διαδικασία μάθησης, η οποία θα επιβλέπεται από έναν ειδικό, ώστε να μπορέσουν να χρησιμοποιηθούν στην επίλυση προβλημάτων. Βασικό εμπόδιο στην καθολική αποδοχή των τεχνητών νευρωνικών δικτύων είναι η αδιαφάνεια που περιβάλλει τη διαδικασία λήψης μιας απόφασης, η ανικανότητά τους δηλαδή να εξηγήσουν στο χρήστη με τρόπο κατανοητό, πώς απέκτησαν τη συγκεκριμένη γνώση, πώς οδηγήθηκαν στη συγκεκριμένη απόφαση, ποιες οι συνθήκες που επικρατούν και πώς εκτελούνται οι διάφορες διεργασίες στο εσωτερικό τους.

Το 1965, ο L.Zadeh, θεμελίωσε ένα νέο τρόπο περιγραφής και αποτύπωσης της γνώσης με τη θεωρία του περί «ασαφών συνόλων» [28]. Πίσω από τη θεωρία αυτή, που είναι γνωστή και ως «ασαφής λογική», βρίσκεται η ίδια η ανθρώπινη συμπεριφορά, καθώς η «ασάφεια» είναι μια πολύ συνηθισμένη έννοια που εμπεριέχεται σε πολλές ανθρώπινες δραστηριότητες, αλλά και στην περιγραφή πολλών εννοιών και ποσοτήτων. Τα συστήματα απόφασης ή ελέγχου που λειτουργούν σε ένα περιβάλλον αβεβαιότητας και στηρίζονται στη θεωρία της ασαφούς λογικής, λέγονται ασαφή συστήματα και έχουν βρει ποικίλες εφαρμογές στην ιατρική, στον αυτόματο έλεγχο, αλλά και στην κατασκευή σύγχρονων «έξυπνων» ηλεκτρικών συσκευών. Η

γενική αρχιτεκτονική των ασαφών συστημάτων περιλαμβάνει τέσσερις μονάδες επεξεργασίας: μια μονάδα ασαφοποίησης, η οποία μετατρέπει τα δεδομένα εισόδου σε ασαφή σύνολα, μια βάση ασαφών κανόνων της μορφής EAN- TOTE (ασαφής βάση γνώσης), μια ασαφή συλλογιστική μηχανή που εξάγει τα ασαφή συμπεράσματα, και μια μονάδα ασαφοποίησης, η οποία μετατρέπει τα ασαφή συμπεράσματα σε σαφώς καθορισμένης μορφής. Για να κατασκευαστεί όμως ένα ασαφές σύστημα είναι αναγκαία η ύπαρξη ενός εμπειρογνώμονα που θα δημιουργήσει ένα σύνολο ασαφών κανόνων για το συγκεκριμένο πρόβλημα. Λόγω της δυσκολίας αυτής της αναγκαίας συνθήκης, τα ασαφή συστήματα, πολύ σύντομα μετά τη γέννησή τους, αναζήτησαν τρόπους να εξάγουν τη γνώση μέσα από τα δεδομένα του προβλήματος. Η απάντηση σε αυτούς τους προβληματισμούς, ήρθε με τον υβριδικό συνδυασμό των δυο παραπάνω μαθηματικών μοντέλων, που εκμεταλλεύεται τα πλεονεκτήματα των νευρωνικών δικτύων και της ασαφούς συλλογιστικής, και παράλληλα στοχεύει στην εξάλειψη των μειονεκτημάτων τους [10,27]. Έτσι τα νευροασαφή συστήματα διαθέτουν την ικανότητα να παράγουν ασαφή γνώση με κατανοητή μορφή, από τα δεδομένα (κανόνες EAN-TOTE), προσαρμόζοντας τις παραμέτρους τους στο σύστημα και εξαλείφοντας την ανάγκη εισαγωγής έτοιμης γνώσης από ειδικούς. Για την ικανότητά τους αυτή χαρακτηρίζονται προσαρμοστικά.

Τα σύγχρονα προσαρμοστικά νευροασαφή συστήματα απεικονίζονται συνήθως ως πολυστρωματικά δίκτυα προσοτροφοδότησης (multi-layered feed forward networks), όπου τα βάρη των συνάψεων και οι συναρτήσεις ενεργοποίησης διαφέρουν από τα κοινά νευρωνικά δίκτυα.

Για τη συγκεκριμένη διπλωματική εργασία, θεωρούμε ότι ένα προσαρμοστικό νευροασαφές σύστημα διέπεται από τις παρακάτω αρχές:

- Ως νευροασαφές σύστημα ορίζεται το ασαφές σύστημα που χρησιμοποιεί έναν αλγόριθμο μάθησης, ο οποίος προέρχεται ή βασίζεται στη θεωρία των νευρωνικών δικτύων. Προσπαθεί να προσεγγίσει μια άγνωστη συνάρτηση N διαστάσεων, η οποία ορίζεται ως ένα βαθμό από τα δεδομένα.
- Ένα νευροασαφές σύστημα μπορεί να αποδοθεί ως ένα δίκτυο προσοτροφοδότησης με 3 επίπεδα. Το πρώτο επίπεδο αντιπροσωπεύει τις μεταβλητές εισόδου, το μεσαίο (ιρυσό) επίπεδο αντιπροσωπεύει τους ασαφείς κανόνες, και το τρίτο επίπεδο συμβολίζει τους μεταβλητές εξόδου. Τα ασαφή σύνολα (fuzzy sets) κωδικοποιούνται ως ασαφή βάρη συνάψεων χρησιμοποιώντας μια από τις γνωστές συναρτήσεις συμμετοχής (π.χ τριγωνική, τραπεζοειδής, γκαουσιανή κλπ.)
- Ένα νευροασαφές σύστημα είναι ικανό κάθε στιγμή (δηλαδή πριν, κατά τη διάρκεια και μετά την εκπαίδευση), να μεταφραστεί ως ένα σύνολο ασαφών κανόνων, με δομή και παραμέτρους που έχουν φυσική και νοηματική υπόσταση. Επιπλέον είναι δυνατό να δημιουργηθεί ένα σύστημα απευθείας από τα δεδομένα, ή να

αρχικοποιηθεί με προγενέστερη γνώση, η οποία να έχει τη μορφή ασαφών κανόνων «εάν-τότε».

- Ένα νευροασαφές σύστημα επιδεικνύει την ικανότητα να προσαρμόζεται τόσο σε αριθμητικά όσο και σε γλωσσικά περιβάλλοντα.

Πληθώρα τέτοιων συνεργαζόμενων μοντέλων έχουν προταθεί στη βιβλιογραφία και περιλαμβάνουν περιπτώσεις για: προσεγγιστική συλλογιστική, εξαγωγή συμπερασμάτων και έλεγχο [9, 12, 15], ταξινόμηση [3, 18], παραγωγή γνώσης από αριθμητικά δεδομένα [16] και απλοποίηση ή περιορισμό των κανόνων [11, 19]. Σε γενικές γραμμές, οι μέθοδοι ανάπτυξης προσαρμοστικών νευροασαφών μοντέλων παρουσιάζουν κάποια κοινά σημεία τα οποία προέρχονται από την επιθυμία:

1. Να διοχετεύσουμε γνώση οδηγούμενη από δεδομένα σε μια αρχιτεκτονική δικτύου για την υλοποίηση ταχείας εκμάθησης.
2. Να σχεδιάσουμε τον κατάλληλο μηχανισμό σύνθεσης και συγκέντρωσης στοιχείων που να μπορεί ταυτόχρονα να χειρίζεται αριθμητικά και γλωσσικά χαρακτηριστικά, έτσι ώστε να παράγει ορθά συμπεράσματα.
3. Να συμπεριλάβουμε μια διαδικασία προσαρμογής των κανόνων μέσω της εκπαίδευσης από τα δεδομένα.
4. Να εξάγουμε και να μεταφράζουμε την εκπαιδευμένη γνώση σαν μια βάση κατανοητών κανόνων.

1.3 Περιγραφή Διπλωματικής

1.3.1 Σκοπός – Κίνητρα

Όπως προκύπτει από την περιγραφή των νευροασαφών συστημάτων, πρόκειται για συστήματα που μπορούν να ικανοποιήσουν ένα ευρύ φάσμα αναγκών της καθημερινής ζωής, οι οποίες όσο περνάνε τα χρόνια γίνονται όλο και περισσότερες και πιο πολύπλοκες. Και αυτό γιατί στις καθημερινές ανθρώπινες ανάγκες, έρχονται να προστεθούν και διάφοροι περιβαλλοντικοί, οικονομικοί και κοινωνικοί παράγοντες που δημιουργούν νέες απαιτήσεις, όχι μόνο για τα νευροασαφή συστήματα, αλλά και για κάθε τομέα της τεχνολογίας.

Στις μέρες μας οι ανάγκες για οικονομία, όχι μόνο ενεργειακή αλλά και οικονομία πόρων, χρόνου και χρήματος, δεν αφήνουν περιθώρια για περιττές λειτουργίες των συστημάτων και αδικαιολόγητες δαπάνες. Το ζητούμενο είναι η καλύτερη και πιο αποτελεσματική λειτουργία τους, με το μικρότερο ενεργειακό και οικονομικό κόστος, με τους ελάχιστους δυνατούς πόρους και στο μικρότερο δυνατό χρόνο. Επομένως θέλουμε συστήματα με μια ευέλικτη δομή που θα διαχειρίζονται με σωστό και αποτελεσματικό τρόπο τους παρεχόμενους πόρους, ελαχιστοποιώντας τις απώλειες και μεγιστοποιώντας την απόδοσή τους, επιθυμία

που εξυπηρετεί την εξοικονόμηση χρόνου και ενέργειας, άρα και χρημάτων, στόχους που όλοι πλέον προσπαθούν να πετύχουν.

Η πλειοψηφία των μελετών για την πλέον οικονομική λειτουργία των συστημάτων, ασχολείται κυρίως με τη διαχείριση και επεξεργασία των παραμέτρων που εισάγονται σε αυτά τα συστήματα, ώστε πριν ακόμα τεθεί ένα σύστημα σε λειτουργία να έχουμε εξασφαλίσει την καλύτερη και οικονομικότερη χρήση του.. Η προετοιμασία των δεδομένων εισόδου και η προεπεξεργασία τους, είναι από τα βασικά αντικείμενα έρευνας προς την επίτευξη μιας «οικονομικότερης» λειτουργίας των συστημάτων. Η αποβολή της περιττής πληροφορίας στην είσοδο του συστήματος, και η προσαρμογή των παραμέτρων του, σίγουρα αποτελούν βήματα προς αυτήν την κατεύθυνση. Πρόκειται συνήθως για μεθόδους που δεν απαιτούν πολύ χρόνο και δεν είναι δύσκολο να εφαρμοστούν στην πράξη.

Λίγες όμως είναι οι προσπάθειες που έχουν γίνει για τη μεταβολή της ίδιας της εσωτερικής δομής ενός συστήματος. Οι περισσότερες μεθοδολογίες αναφέρονται σε μεταβολή παραμέτρων αλλά όχι σε μεταβολή της εσωτερικής δομής. Κάτι τέτοιο βέβαια είναι και λογικό μέχρι ένα βαθμό, αφού η εσωτερική δομή ενός συστήματος, και ιδιαίτερα ενός νευροασαφούς συστήματος, δεν είναι πάντα εύκολα κατανοητή. Ο τρόπος που επεξεργάζεται ένα τέτοιο σύστημα τα δεδομένα και παίρνει τις αποφάσεις, παραμένει συγκεχυμένος και πολύπλοκος. Κατά συνέπεια η εσωτερική δομή ενός τέτοιου συστήματος δεν επιδέχεται εύκολα επεξεργασία και μεταβολές. Ένα σύστημα όμως που θα ελίσσεται και θα προσαρμόζει την εσωτερική του δομή στις εκάστοτε απαιτήσεις και τις τρέχουσες συνθήκες κάθε προβλήματος, προσφέρει σίγουρα περισσότερες δυνατότητες εφαρμογής και λειτουργεί ακόμα πιο «ανθρώπινα», εξυπηρετώντας έτσι καλύτερα τους σκοπούς μας. Επιπλέον, αν και η αποκοπή «περιττής» πληροφορίας από τις εισόδους του συστήματος είναι μια κλασσική μέθοδος μεταβολής παραμέτρων, πάντα είναι προτιμότερο να μη χάνεται καθόλου πληροφορία, αφού και η παραμικρή απώλεια σίγουρα επηρεάζει έστω και ελάχιστα το αποτέλεσμα του συστήματός μας. Με τη μεταβολή στη δομή του συστήματος κάτι τέτοιο αποφεύγεται.

Λαμβάνοντας υπόψη όλες αυτές τις απαιτήσεις που πρέπει να ικανοποιεί ένα σύστημα, καταλήγουμε σε ένα συμπέρασμα για τη μέθοδο σχεδιασμού που είναι προτιμότερο να χρησιμοποιήσουμε. Όπως αναφέρθηκε, τα νευροασαφή συστήματα παρέχουν τη δυνατότητα επεξεργασίας ασαφούς- μη ακριβούς πληροφορίας. Η ασάφεια είναι ένα από τα βασικά χαρακτηριστικά που εμφανίζεται στον ανθρώπινο τρόπο σκέψης, με την έννοια ότι σε πολλές περιπτώσεις ο άνθρωπος δεν αντιμετωπίζει τις καταστάσεις με τρόπο απόλυτο και μονοσήμαντο. Η δυνατότητα επομένως χειρισμού ασαφούς πληροφορίας που παρέχουν τα νευροασαφή συστήματα, είναι απαραίτητη για να πετύχουμε την επιθυμητή προσαρμοστικότητα ενός ευφυούς συστήματος, το ένα δηλαδή από τα δυο αναγκαία χαρακτηριστικά του.

Περιβαλλοντικοί, οικολογικοί και πρακτικοί λόγοι μας ωθούν όμως να αναζητήσουμε επιπλέον τρόπους βελτίωσης στη λειτουργία των συστημάτων αυτών, ώστε να εμφανίζονται και

το δεύτερο αναγκαίο χαρακτηριστικό, αυτό της «οικονομικότερης» λειτουργίας. Η ανάγκη για ενεργειακή οικονομία, σωστή κατανάλωση και οικονομία πόρων, μας οδηγεί στην αναζήτηση νέων μεθόδων διαχείρισης της δομής και των παραμέτρων ενός νευροασαφούς συστήματος, πέρα από την προεπεξεργασία των δεδομένων εισόδου και την αποκοπή της «περιττής» πληροφορίας που κυρίως εφαρμόζονται μέχρι τώρα.

Στα πλαίσια της έρευνας για τη σωστή διαχείριση των πόρων και της δομής ενός συστήματος κινείται και η παρούσα διπλωματική εργασία. Έχοντας ως κίνητρο τη βελτίωση των νευροασαφών συστημάτων, ιδιαίτερα στον τομέα της πιο «οικονομικής» και φιλικής προς το περιβάλλον λειτουργίας τους, συνδυάσαμε συστήματα και μεθόδους, και μέσα από τη σύγκριση πειραματικών αποτελεσμάτων καταλήξαμε σε συμπεράσματα για το σύστημα που ανταποκρίνεται περισσότερο στις ανάγκες μας.

Σκοπός είναι να καταλήξουμε σε ένα σύστημα το οποίο θα εμφανίζει μεγάλη προσαρμοστικότητα και ευελιξία, θα μπορεί να χρησιμοποιηθεί σε ένα ευρύ φάσμα εφαρμογών, θα παρέχει δηλαδή τη δυνατότητα της γενίκευσης, και ταυτόχρονα θα ικανοποιεί και τις ανάγκες για μικρό κόστος σε όλους τους τομείς.

Στην παρούσα διπλωματική χρησιμοποιήσαμε ένα νευροασαφές σύστημα γινομένου, το ARANFIS (Adaptive Resource Allocating Neural Fuzzy System), το οποίο αναλύεται στη συνέχεια της διπλωματικής, στο οποίο εφαρμόσαμε τεχνικές μεταβολής τόσο των παραμέτρων όσο και της εσωτερικής δομής του. Η δοκιμή του συστήματος έγινε κυρίως σε κλασσικά προβλήματα ταξινόμησης. Πρόκειται για ένα μοντέλο συστήματος προσαρμοζόμενης δομής με ανάθεση πόρων. Χρησιμοποιήσαμε την ανάλυση ευαισθησίας, ένα εργαλείο στατιστικής μελέτης που χρησιμοποιείται ευρέως και στο χώρο των ευφών συστημάτων. Η δική μας προσέγγιση περιλαμβάνει την εφαρμογή της ανάλυσης ευαισθησίας όχι όμως σε επίπεδο των δεδομένων εισόδου του συστήματος, όπως έχει γίνει στις μέχρι τώρα μελέτες, αλλά στο επίπεδο των κανόνων. Επιπλέον για την αξιολόγηση και την ιεραρχική επιλογή των κανόνων εφαρμόσαμε μια νέα μέθοδο σύμφωνα με την οποία το σύνολο των κανόνων χωρίζεται μόνο σε δυο υποσύνολα, το κύριο και το δευτερεύον. Σκοπός της έρευνάς μας η καλύτερη δυνατή επίδοση του συστήματος με τη χρήση των λιγότερων δυνατών πόρων.

1.3.2 Διάρθρωση Διπλωματικής

Η παρούσα διπλωματική εργασία χωρίζεται σε 5 κεφάλαια. Στο παρόν κεφάλαιο, που είναι και το πρώτο της διπλωματικής, γίνεται μια παρουσίαση της σύγχρονης επιστημονικής πραγματικότητας στο χώρο των ευφών και των νευροασαφών συστημάτων, και γίνεται μια προσπάθεια να εντοπιστούν τα πεδία που επιδέχονται περαιτέρω έρευνα και βελτίωση. Από εκεί προκύπτει και ο στόχος που εξυπηρετεί αυτή η διπλωματική εργασία. Στο κεφάλαιο 2 περιγράφονται συνοπτικά τα νευροασαφή συστήματα και η ανάλυση ευαισθησίας, δυο

επιστημονικά πεδία με τα οποία ασχολείται εκτενώς η διπλωματική, και τα οποία συνδυάζονται για να χρησιμοποιηθούν στην έρευνά μας. Στο κεφάλαιο 3 αναλύεται η μεθοδολογία που ακολουθήσαμε. Δίνεται μια περιγραφή των συγκεκριμένων νευροασφαρών συστημάτων που θα χρησιμοποιήσουμε στη διπλωματική (Supfunis, Aranfis) που συνοδεύεται από μια σύγκριση μεταξύ τους, ώστε να γίνει σαφές σε ποια επιπλέον στοιχεία που διακρίνουν το Aranfis στηρίχτηκε η έρευνά μας. Επιπλέον αναλύεται ο τρόπος με τον οποίο χρησιμοποιήθηκε η ανάλυση ευαισθησίας στη μελέτη μας και ο συνδυασμός νευροασφαρών συστημάτων και ανάλυσης ευαισθησίας, ενώ περιγράφεται και η μέθοδος που χρησιμοποιήθηκε για την αξιολόγηση, ιεράρχηση και δυναμική επιλογή των κανόνων. Το κεφάλαιο 4 περιλαμβάνει τα πειραματικά αποτελέσματα που προέκυψαν, ενώ το κεφάλαιο 5, το τελευταίο κεφάλαιο της διπλωματικής, περιλαμβάνει συνολικά συμπεράσματα και πορίσματα της έρευνας, καθώς και πιθανές μελλοντικές επεκτάσεις. Ακολουθεί η βιβλιογραφία στην οποία στηρίχτηκε η παρούσα διπλωματική, καθώς και το βιογραφικό της συγγραφέως.

Κεφάλαιο 2:Νευροασαφή συστήματα γινόμενου – Ανάλυση Ευαισθησίας

2.1 Γενικά

Στο κεφάλαιο αυτό παρουσιάζονται αναλυτικά τα επιστημονικά πεδία που αποτελούν τη βάση αυτής της διπλωματικής. Αρχικά θα παρουσιάσουμε γενικά τα νευροασαφή συστήματα γινόμενου ώστε να γίνει κατανοητός ο λόγος που επιλέξαμε να τα χρησιμοποιήσουμε. Στη συνέχεια θα δοθεί και μια περιγραφή της ανάλυσης ευαισθησίας η οποία χρησιμοποιήθηκε κατά την έρευνά μας και είναι ένα από τα πιο χρήσιμα στατιστικά εργαλεία με ποικίλες εφαρμογές σε πολλούς τομείς και πέρα από αυτόν που εξετάζεται σε αυτήν τη διπλωματική.

2.2 Μοντελοποίηση νευροασαφών συστημάτων και προσδιορισμός δομής.

Η μοντελοποίηση νευροασαφών συστημάτων αποτελεί ένα σημαντικό τμήμα του πεδίου της υβριδικής υπολογιστικής νοημοσύνης. Σκοπός της είναι η κατασκευή ενός μηχανισμού λήψης αποφάσεων που είναι ικανός να προβλέπει και να εξηγεί τη συμπεριφορά ενός άγνωστου συστήματος, το οποίο περιγράφεται με ένα σύνολο από δεδομένα. Ο προσδιορισμός ενός νευροασαφούς συστήματος, αποτελείται βασικά από δύο κύρια στάδια: το στάδιο του προσδιορισμού δομής και το στάδιο του προσδιορισμού παραμέτρων [26]. Ο προσδιορισμός δομής (Structure Identification) περιλαμβάνει τον καθορισμό των μεταβλητών εισόδου και εξόδου, του αριθμού κανόνων στη βάση γνώσης, αλλά και τον ορισμό των σχέσεων και συνδέσεων μεταξύ όλων των μεταβλητών του συστήματος. Επιπλέον, μέσω του προσδιορισμού δομής, μπορεί να προκύψει μια αρχική εκτίμηση των συναρτήσεων συμμετοχής που συνιστούν τα βάρη εισόδου και εξόδου, το οποίο συμβάλλει όχι μόνο στην καλύτερη απόδοση του συστήματος, αλλά και στην επιτάχυνση της εκπαίδευσής του [10]. Ο προσδιορισμός παραμέτρων από την άλλη (Parameter Identification) σχετίζεται με τη ρύθμιση, συνήθως μέσω εκπαίδευσης, των ελεύθερων παραμέτρων του συστήματος, όπως είναι οι συναρτήσεις συμμετοχής και τα βάρη εισόδου και εξόδου (antecedent- consequent weights).

Επικρατέστερη τεχνική σήμερα για τη ρύθμιση των παραμέτρων ενός προσαρμοστικού νευροασαφούς μοντέλου είναι αυτή της επιβλεπόμενης μάθησης με τη χρήση της μεθόδου αναστροφής διάδοσης (back propagation) που βασίζεται στην ελαχιστοποίηση του τετραγώνου του σφάλματος, ενώ ως εναλλακτικές λύσεις εμφανίζονται ο κανόνας Widrow-Hoff και το φίλτρο Kalman [26]. Η επιλογή γενετικών αλγορίθμων για τη ρύθμιση των παραμέτρων αποτελεί επίσης μια αρκετά διαδεδομένη, αλλά και αρκετά χρονοβόρα τεχνική.

Για τον προσδιορισμό δομής δεν υφίσταται κάποια γενική και κυρίαρχη μεθοδολογία. Τα πιο σημαντικά από τα προβλήματα που παρουσιάζονται είναι η εύρεση του βέλτιστου αριθμού κανόνων και η «κατάρρα της διάστασης» (curse of dimensionality). Το τελευταίο αναφέρεται στο φαινόμενο κατά το οποίο, όσο αυξάνεται γραμμικά η διάσταση του προβλήματος, τόσο αυξάνεται εκθετικά ο αριθμός των κρυφών κόμβων του δικτύου που απαιτούνται για την επίλυσή του.

Νευροασαφή συστήματα γινομένου

Τα νευροασαφή συστήματα γινομένου είναι ασαφή μοντέλα που χρησιμοποιούν το αλγεβρικό γινόμενο ως τελεστή συσσώρευσης (aggregation operator) των δράσεων σε έναν κόμβο κανόνα, αντί του τελεστή ελαχίστου (min) που συναντάται πιο συχνά στη βιβλιογραφία [12]. Ο τελεστής συσσώρευσης αποτελεί τη διεργασία συλλογισμού

(implication process) ενός κανόνα EAN-TOTE, ή αλλιώς τη συνάρτηση ενεργοποίησής του (activation function).

Η χρήση του γινομένου ως συνάρτηση ενεργοποίησης παρουσιάζει συγκεκριμένα πλεονεκτήματα. Το βασικότερο από αυτά, είναι πως η συνάρτηση γινομένου είναι συνεχής, διαφορίσιμη, συμμετρική και ομαλή, στοιχεία που επιτρέπουν την εκπαίδευση του δικτύου με χρήση του αλγορίθμου ανάστροφης διάδοσης για τη ρύθμιση των παραμέτρων του συστήματος. Αντίθετα με τον πιο συνηθισμένο τελεστή ελαχίστου, το γινόμενο δεν αγνοεί την πληροφορία που αφορά στη διάσταση του προβλήματος και μπορεί να εκτιμήσει καλύτερα την ισχύ των χαρακτηριστικών εισόδου. Επίσης, για ένα μεγάλο εύρος διασποράς, ο τελεστής γινομένου, είναι ικανός να διακρίνει τις εισόδους που είναι όμοιες με το διάνυσμα βαρών και αυτές που είναι ανόμοιες. Με άλλα λόγια, ο τελεστής γινομένου εμφανίζεται ως πιο ευαίσθητος στις διαφορές των προτύπων από άλλους τελεστές.

Η συνάρτηση γινομένου χαρακτηρίζεται και από ένα μειονέκτημα. Δεδομένου πως το γινόμενο αποτελείται από παράγοντες των οποίων το πλήθος είναι όσο η διάσταση του προβλήματος και θεωρώντας πως ο κάθε παράγοντας λαμβάνει, λόγω κανονικοποίησης, τιμές μεταξύ μηδέν και ένα σε προβλήματα μεγάλης διάστασης το γινόμενο παράγει τιμές πολύ κοντά στο μηδέν. Θεωρητικά, δεδομένου ότι αναφερόμαστε σε υπολογιστική νοημοσύνη, υπάρχει ένας πεπερασμένος αριθμός παραγόντων πάνω από τον οποίο, το γινόμενο σε μια υπολογιστική μηχανή μηδενίζεται αφού η υποδιαστολή υπερχειλίζει. Επομένως τα νευροασαφή συστήματα γινομένου έχουν ένα άνω όριο διάστασης που μπορούν να χειριστούν με ασφάλεια. Μια λύση σε αυτό το φαινόμενο είναι η κανονικοποίηση των τιμών ενεργοποίησης ως προς το άθροισμά τους.

Ανάλυση ευαισθησίας

Στις περισσότερες πραγματικές εφαρμογές, είναι δύσκολο να διακρίνουμε όλες τις μεταβλητές εισόδου με σημαντική επίδραση στην έξοδο του συστήματος. Ουσιαστικά η ανάλυση ευαισθησίας είναι ένας τρόπος για να διακρίνουμε την επίδραση που έχει πάνω σε μια μεταβλητή X , η οποιαδήποτε μεταβολή που υφίσταται μια άλλη μεταβλητή Y [4, 6]. Μπορούμε επομένως να αποφανθούμε για το ποια μεταβλητή επηρεάζει σε μεγάλο βαθμό τα αποτελέσματα του συστήματος. Αυτό μας βοηθάει στο να επικεντρωθούμε σε μικρότερο όγκο δεδομένων και να μελετήσουμε περισσότερες περιπτώσεις για μια μεταβλητή που ξέρουμε ότι οι μεταβολές στην τιμή της θα έχουν σημαντική επίδραση στο σύστημα και τα δεδομένα μας. Μεταβλητές οι οποίες, όπως προκύπτει από την ανάλυση ευαισθησίας που κάνουμε, όταν μεταβάλλουν την τιμή τους δεν επηρεάζουν ιδιαίτερα το αποτέλεσμά μας, θα μας απασχολήσουν λιγότερο στη μελέτη μας, ίσως και καθόλου.

Υπάρχουν διάφοροι τρόποι να εφαρμόσουμε την ανάλυση ευαισθησίας στα δεδομένα μας, ανάλογα με το μοντέλο που εξετάζουμε. Στις περισσότερες όμως των περιπτώσεων εξετάζουμε ξεχωριστά την επίδραση της μεταβολής κάθε μεταβλητής μεμονωμένα, και στη

συνέχεια καταλήγουμε σε κάποιο συμπέρασμα για το ποια συνδέεται περισσότερο με το αποτέλεσμα μας. Επομένως η ευαισθησία της εξόδου ενός νευρωνικού- νευροασαφούς δικτύου ως προς την εισόδό του, είναι ένα μέτρο εκτίμησης της απόδοσης του δικτύου, εξίσου σημαντικό με το σφάλμα και την ικανότητα γενίκευσης[29].

Στο πεδίο των νευρωνικών δικτύων που μας ενδιαφέρει εδώ, η ανάλυση ευαισθησίας έχει εφαρμοστεί σε αρκετές ερευνητικές μελέτες, με ζητούμενο την αξιολόγηση της αλληλεπίδρασης μεταξύ των κόμβων εισόδου και εξόδου, ή με άλλα λόγια η ανεύρεση των χαρακτηριστικών του προβλήματος που συμβάλλουν πιο πολύ στην αποτελεσματικότητα του συστήματος [6, 7, 30]. Προφανώς, εκείνες οι εισοδοί που επηρεάζουν λιγότερο, μπορούν να θεωρηθούν ως λιγότερο σημαντικές για την έξοδο και να απομακρυνθούν από το σύστημα χωρίς αυτό να επιφέρει ιδιαίτερη μείωση στην απόδοση. Το μέτρο ευαισθησίας σε αυτές τις περιπτώσεις υπολογίζεται από τον Ιακωβιανό πίνακα των μερικών παραγώγων της εξόδου θ_{ψ_k} ως προς τις μερικές παραγώγους της εισόδου θ_{x_i} , δηλαδή $\theta_{\psi_k}/\theta_{x_i}$. Βέβαια η ανάλυση αυτή προϋποθέτει ότι όλες οι μεταβλητές εισόδου είναι αριθμητικές και συνεχείς. Σε αντίθετη περίπτωση, όταν οι μεταβλητές λαμβάνουν, είτε συμβολικές- λεκτικές τιμές, είτε διακριτές, ο υπολογισμός των μερικών παραγώγων δεν είναι εφικτός. Με την κλασική επομένως εφαρμογή της ανάλυσης ευαισθησίας στον τομέα των νευρωνικών δικτύων, συνδέουμε την εισοδο ενός συστήματος, μόνο όταν πρόκειται για συνεχείς αριθμητικές μεταβλητές, με την έξοδο τους, σε μια προσπάθεια να αγνοήσουμε τις μεταβλητές εισόδου που δεν επηρεάζουν σε μεγάλο βαθμό την έξοδό μας, ώστε να μειώσουμε τον όγκο των δεδομένων προς επεξεργασία.

Καθώς η κλασική εφαρμογή της ανάλυσης ευαισθησίας μας περιορίζει ως προς το είδος της εισόδου, προκύπτει η ανάγκη για βελτίωση στον τομέα αυτό.

Μια νέα πρόταση στο πεδίο των νευροασαφών συστημάτων γινομένου προωθεί την εφαρμογή της ανάλυσης ευαισθησίας στο επίπεδο των κανόνων και όχι σε αυτό της εισόδου του δικτύου. Για να υλοποιηθεί αυτό θεωρείται ως μεταβλητή επιρροής της εξόδου η ισχύς ενεργοποίησης z_j του κανόνα j . Η επιλογή αυτή υποστηρίζεται από δυο σημαντικά επιχειρήματα. Πρώτον, η ισχύς ενεργοποίησης είναι μια εξ ορισμού ποσοτική μεταβλητή και άρα δεν μπορεί να φέρει συμβολικές ή λεκτικές τιμές. Δεύτερον, οι πιο διαδεδομένες συναρτήσεις ενεργοποίησης σε νευροασαφή συστήματα είναι ομαλές, συνεχείς και διαφορίσιμες. Επομένως μπορούμε να εφαρμόσουμε τη μέθοδο αυτή σε ένα ευρύ φάσμα αλγορίθμων, ανεξάρτητα από τη φύση των δεδομένων και του προβλήματος.

2.3 Κίνητρο και Συμβολή

Σε πολλά νευροασαφή συστήματα που περιγράφονται σε διεθνή συνέδρια και επιστημονικά περιοδικά, η διαδικασία μοντελοποίησης αρκείται στον προσδιορισμό των παραμέτρων, καθώς το πλήθος των απαιτούμενων κανόνων επιλέγεται εμπειρικά και μένει σταθερό κατά την εκπαίδευση του δικτύου, ενώ η αρχικοποίηση του συστήματος πραγματοποιείται με τυχαίες τιμές. Η τυχαία και στατική αυτή δομή περιορίζει την επιθυμητή ιδιότητα της προσαρμογής σε νέες καταστάσεις, άλλα και της ικανότητας γενίκευσης. Σε άλλες περιπτώσεις, όταν χρησιμοποιείται προσδιορισμός δομής, αυτός εφαρμόζεται μερικώς, κυρίως για την αρχικοποίηση των παραμέτρων, ώστε να αποφευχθεί όσο είναι δυνατόν, το φαινόμενο παγίδευσης του αλγορίθμου μάθησης σε τοπικό ελάχιστο. Η αρχικοποίηση τότε καθορίζεται συνήθως από κάποιο αλγόριθμο ομαδοποίησης ή διαχωρισμού και αναζητά έναν αυθαίρετο αριθμό κανόνων που θα επιλύσουν το πρόβλημα. Αυτές οι εμπειρικές μέθοδοι έχουν συνήθως εμπειρική εφαρμογή και απαιτούν αρκετές παραμέτρους για να λειτουργήσουν.

Η πλήρης διαδικασία προσδιορισμού δομής σε όλα τα επίπεδα εμφανίζεται σε ελάχιστες περιπτώσεις στη βιβλιογραφία. Όταν αυτό συμβαίνει, ο προσδιορισμός δομής αποτελεί προστάδιο της εκπαίδευσης, κατά τη διάρκεια της οποίας η δομή του νευροασαφούς συστήματος παραμένει αμετάβλητη, με αποτέλεσμα την περιορισμένη ευκαμψία και προσαρμοστικότητα του συστήματος. Το φαινόμενο της στατικής δομής είναι πολύ πιο έντονο σε καταστάσεις που το σύστημα βρίσκεται σε λειτουργία. Αυτό δικαιολογείται, αφού η προσαρμογή σε ένα πρόβλημα είναι πολύ πιο δύσκολη όταν δεν υπάρχουν οι επιθυμητές τιμές εξόδου, ώστε να καταστεί εφικτή κάποια τεχνική επιβλεπόμενης μάθησης. Οι μοναδικές περιπτώσεις που έχουμε μεταβλητή δομή δικτύου, αφορούν στην κατάργηση κάποιων παραμέτρων του δικτύου, συνήθως κόμβων/κανόνων, προτύπων εισόδου, γεγονός που βέβαια συνεπάγεται μερική απόρριψη πληροφορίας που ίσως να φαινόταν χρήσιμη σε κάποιες περιπτώσεις. Ακόμα και παράμετροι εισόδου με πολύ μικρή επίδραση στην έξοδο του συστήματος, είναι πολύ πιθανόν να χρειάζεται να συμπεριληφθούν στη λειτουργία του δικτύου προκειμένου να καλύψουν τις απαιτήσεις για την εξαγωγή του κατάλληλου αριθμού κανόνων. Αυτό βέβαια δεν είναι ευδιάκριτο, αφού η διαδικασία της εξαγωγής κανόνων είναι πολύπλοκη και απρόσιτη για το χρήστη. Για το λόγο αυτό είναι αναγκαία η κατανόηση της εσωτερικής δομής ενός δικτύου, ώστε να γίνει δυνατή και η μεταβολή στη δομή του. [7].

Αν συνοψολογίσουμε και την αργή επεξεργασία των δεδομένων εκπαίδευσης και τη δυσκολία χειρισμού πολλών διαστάσεων εισόδου («curse of dimensionality»), στοιχεία που συχνά συναντάμε στα νευροασαφή συστήματα, εύκολα αντιλαμβάνεται κανείς πως ο χώρος προσφέρεται για έρευνα και παρουσιάζει αρκετά προβλήματα προς επίλυση.

Όσον αφορά στην ανάλυση ευαισθησίας, όπως αναφέρθηκε και παραπάνω, η καινοτομία που προτείνεται είναι η εφαρμογή της μεθόδου στο επίπεδο των κόμβων κανόνων σε σχέση με την έξοδο του συστήματος και όχι στο επίπεδο των δεδομένων εισόδου, όπως εφαρμοζόταν μέχρι σήμερα. Με αυτόν τον τρόπο επεμβαίνουμε και πάλι στην εσωτερική δομή του

| Κεφάλαιο 2: Νευροασαφή συστήματα γινόμενου – Ανάλυση Ευαισθησίας

δικτύου και δοκιμάζουμε να τη μεταβάλλουμε κατά τη διάρκεια της εκπαίδευσής του. Η ανάλυση ευαισθησίας είναι ένα πολύ συνηθισμένο στατιστικό και ερευνητικό εργαλείο με απλή, σχετικά, μαθηματική εφαρμογή, αφού δεν απαιτεί παρά τον υπολογισμό μιας μερικής παραγώγου. Η εφαρμογή της μεθόδου σε ένα καινούριο κομμάτι του συστήματος μπορεί να έχει σημαντικά αποτελέσματα στην απόδοσή του, χωρίς να επιβαρύνει το υπολογιστικό και σχεδιαστικό του κόστος, είτε το κόστος λειτουργίας του.

Κεφάλαιο 3: Μεθοδολογία- Περιγραφή Τεχνικής

3.1 Γενικά

Στο παρόν κεφάλαιο παρουσιάζονται λεπτομερώς τα συγκεκριμένα νευροασαφή συστήματα που χρησιμοποιήθηκαν στην παρούσα διπλωματική, και ερμηνεύεται ο ρόλος της ανάλυσης ευαισθησίας σε αυτήν τη μελέτη. Δίνεται, δηλαδή, μια αναλυτική περιγραφή της μεθοδολογίας που ακολουθήθηκε, της μεθόδου που εφαρμόστηκε για την επιλογή και αξιολόγηση των κανόνων στο συγκεκριμένο νευροασαφές σύστημα και εξηγείται η λογική που οδηγεί στην επιλογή αυτής της μεθόδου, καθώς και τα πλεονεκτήματά της έναντι άλλων.

3.2 Supfunis

Ένα από τα πιο σύγχρονα και ευέλικτα μοντέλα νευροασαφών συστημάτων γινόμενου είναι το SuPFuNIS[20]. Το μοντέλο SuPFuNIS (Subsethood- Product Fuzzy Neural Inference System) είναι ένα σύστημα γενικής προσέγγισης, το οποίο χρησιμοποιεί την τυπική αρχιτεκτονική των νευροασαφών συστημάτων και μπορεί να περιγραφεί ως ένα δίκτυο προσοτροφόδωσης πλήρους διασύνδεσης με τρία επίπεδα. Ασαφοποιεί όλες τις παραμέτρους του με χρήση συμμετρικών γκαουσιανών συναρτήσεων συμμετοχής και υπολογίζει το βαθμό ενεργοποίησης κάθε κανόνα από το γινόμενο ενός μέτρου ομοιότητας που ονομάζεται αμοιβαία σχέση εγκλεισμού. Αν αντικαταστήσουμε τις συναρτήσεις συμμετοχής με μη συμμετρικές γκαουσιανές κατανομές, τότε αναφερόμαστε στο μη συμμετρικό ASuPFuNIS [24], ενώ η χρήση γενετικών αλγορίθμων για εκπαίδευση του δικτύου παραπέμπει στην «εξελικτική» παραλλαγή του[25].

Το SuPFuNIS επιτυγχάνει υψηλές επιδόσεις και οικονομία σε παραμέτρους, που είναι ένας από τους βασικούς μας στόχους. Αυτό οφείλεται στον πρωτότυπο συνδυασμό τεσσάρων ιδιοτήτων:

1. Χρησιμοποιεί ρυθμιζόμενους ασαφοποιητές εισόδου, οι οποίοι ασαφοποιούν τα εισερχόμενα αριθμητικά δεδομένα εφαρμόζοντας μια γκαουσιανή κατανομή σε κάθε χαρακτηριστικό.
2. Όλη η πληροφορία που κινείται στο δίκτυο είναι σε ασαφή μορφή. Το μοντέλο, επομένως, χρησιμοποιεί ένα μηχανισμό σύνθεσης που στηρίζεται σε ένα μέτρο αμοιβαίας επικάλυψης και ορίζει τη δραστηριότητα που προωθείται σε κάθε κανόνα.
3. Η συνάθροιση των δραστηριοτήτων σε έναν κόμβο κανόνα πραγματοποιείται με τη χρήση του ασαφούς γινόμενου, που έχει ως παράγοντες τις αμοιβαίες επικαλύψεις, ή αλλιώς αμοιβαίες σχέσεις εγκλεισμού.
4. Η έξοδος υπολογίζεται με τη χρήση χωρικής ασαφοποίησης, μέθοδος που είναι παραλλαγή της συνηθισμένης μεθόδου αποσαφοποίησης του κέντρου βάρους.

Η λεπτομερής περιγραφή της δομής και της λειτουργίας του μοντέλου SuPFuNIS δεν κρίνεται εδώ απαραίτητη, αφού το μοντέλο αυτό παρουσιάζει κατάλληλες ιδιότητες και πλεονεκτήματα, αλλά δεν εμφανίζει το στοιχείο της δυναμικής διαχείρισης δομής που μας ενδιαφέρει εδώ. Απλά η δομή και η λειτουργία του αποτελούν μια εισαγωγή για το μοντέλο

που χρησιμοποιήθηκε ως βάση στη διπλωματική, και το οποίο περιγράφεται με λεπτομέρειες παρακάτω.

3.3 Αρχιτεκτονική Τύπου RAN (Resource Allocating Network)

Οι αρχιτεκτονικές νευρωνικών δικτύων με αναδιάταξη πόρων (Resource Allocating Networks –RAN)[22] θεωρούνται κατάλληλες για δυναμική μοντελοποίηση μη στατικών λειτουργιών. Σε αυτήν τη σειριακή μέθοδο εκμάθησης, το δίκτυο αναδιάταξης πόρων αρχικά δεν περιέχει κρυφούς κόμβους. Καθώς εισέρχονται τα δεδομένα εκπαίδευσης, το σύστημα είτε προσθέτει κρυφούς κόμβους, είτε προσαρμόζει τους ήδη υπάρχοντες, ανάλογα με κάποια κριτήρια που βασίζονται συνήθως στο σφάλμα πρόβλεψης. Στη δική μας εκδοχή, εφαρμόζουμε την αρχιτεκτονική RAN [22] σε ένα νευροασαφές σύστημα γινομένου, το οποίο προσαρμόζει και ανανεώνει το πλήθος των κρυφών κόμβων (κανόνων) στις ανάγκες του προβλήματος μέσα από τη διαδικασία της εκπαίδευσης.

Η αρχιτεκτονική τύπου RAN σε ένα δίκτυο περιλαμβάνει το ίδιο το δίκτυο, μια στρατηγική για τον προσδιορισμό καινούριων μονάδων και έναν κανόνα εκπαίδευσης για το δίκτυο. Ένα δίκτυο με αρχιτεκτονική RAN, είναι ένα δίκτυο δυο επιπέδων. Το πρώτο επίπεδο περιλαμβάνει κόμβους που ανταποκρίνονται μόνο σε μια τοπική περιοχή του διαστήματος των δεδομένων εισόδου. Το δεύτερο επίπεδο συσσωρεύει τις εξόδους από αυτούς τους κόμβους και δημιουργεί τη συνάρτηση που προσεγγίζει την απεικόνιση εισόδου-εξόδου για όλο το διάστημα. Οι κόμβοι στο πρώτο επίπεδο εφοδιάζουν μια συγκεκριμένη περιοχή του πεδίου της εισόδου. Όταν η είσοδος ξεφεύγει από τη συγκεκριμένη περιοχή, η απόκριση του κόμβου μειώνεται.

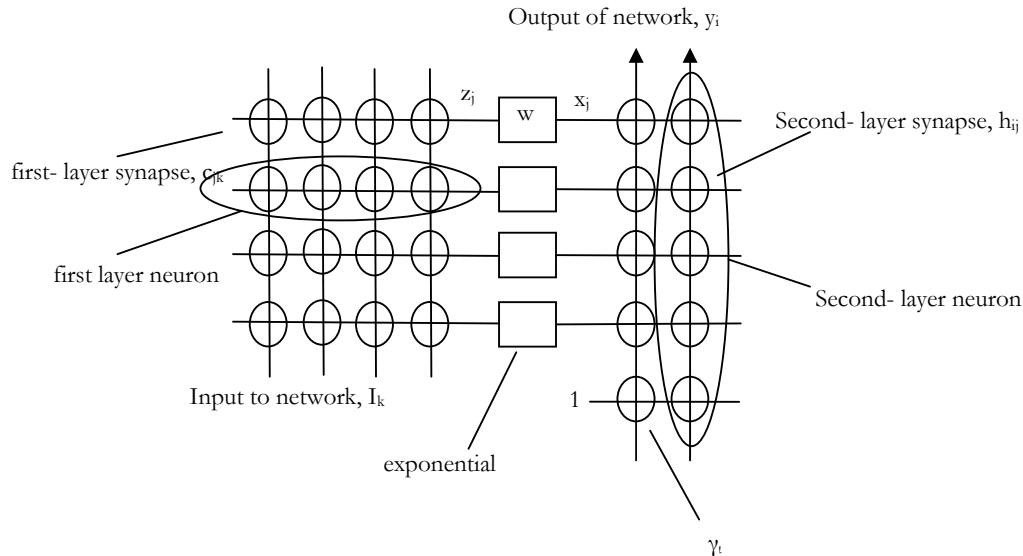
Οι εισοδοί των συνάψεων του δεύτερου επιπέδου, είναι οι εξοδοί των κόμβων του πρώτου επιπέδου. Ο στόχος κάθε σύναψης του δεύτερου επιπέδου, είναι να καθορίσει τη συμβολή κάθε σύναψης του πρώτου επιπέδου σε μια συγκεκριμένη έξοδο του δικτύου. Κάθε έξοδος y του δικτύου είναι το άθροισμα της εξόδου x_i του πρώτου επιπέδου, πολλαπλασιασμένης με το βάρος της αντίστοιχης σύναψης h_i , και μιας διανυσματικής σταθεράς $\vec{\gamma}$, η οποία δεν εξαρτάται από την έξοδο του πρώτου επιπέδου:

$$\vec{y} = \sum_j \vec{h}_j x_j + \vec{\gamma}. \quad 3.1$$

Η σταθερά $\vec{\gamma}$ είναι ουσιαστικά η έξοδος του δικτύου όταν κανένας από τους κόμβους του πρώτου επιπέδου δεν ενεργοποιείται. Το γινόμενο $\vec{h}_j x_j$ μπορεί να θεωρηθεί ως ένας όρος

που προστίθεται ή αφαιρείται από τη σταθερά $\vec{\gamma}$ έτσι ώστε να προκύψει η επιθυμητή συνάρτηση.

Η αρχιτεκτονική τύπου RAN φαίνεται στο Σχήμα 3.1:



Σχήμα 3.1: Παράδειγμα δικτύου με αρχιτεκτονική τύπου RAN

Το δίκτυο του Σχήματος 3.1 υπολογίζει παράλληλα τις αποστάσεις ανάμεσα στο διάνυσμα εισόδου \vec{I} και τα κέντρα c_j .

Αλγόριθμος Εκπαίδευσης

Το δίκτυο ξεκινά αρχικά από μια κενή κατάσταση: δεν έχουν ακόμα φορτωθεί πρότυπα στην είσοδο. Καθώς εισάγονται πρότυπα, το δίκτυο επιλέγει να φορτώσει μερικά από αυτά. Σε κάθε δεδομένο σημείο, το δίκτυο βρίσκεται σε μια συγκεκριμένη κατάσταση, η οποία αντικατοπτρίζει τα πρότυπα που έχουν ήδη φορτωθεί.

Ένας «εντοπιστής» αναγνωρίζει το πρότυπο που δεν αντιπροσωπεύεται σωστά εκείνη τη στιγμή από το δίκτυο και τοποθετεί έναν καινούριο κόμβο που κρατάει το πρότυπο αυτό. Η έξοδος της καινούριας αυτής ενότητας επεκτείνεται στο δεύτερο επίπεδο. Μετά τον εντοπισμό της καινούριας ενότητας, η έξοδος του δικτύου ισούται με την επιθυμητή έξοδο \vec{T} . Οι γραμμικές συνάψεις του δεύτερου επιπέδου του δικτύου γίνονται ίσες με τη διαφορά ανάμεσα στην πραγματική και την επιθυμητή έξοδο του δικτύου.

$$\vec{h}_n = \vec{T} - \vec{y}. \quad 3,2$$

Η αρχιτεκτονική RAN χρησιμοποιεί μια πρωτότυπη συνθήκη με δυο κριτήρια. Ένα ζευγάρι εισόδου –εξόδου (\vec{I}, \vec{T}) θεωρείται ασήμαντο αν:

1. η είσοδος απέχει πολύ από τα κέντρα που ήδη υπάρχουν:

$$\|\vec{I} - \vec{c}_{\text{nearest}}\| > \delta(t), \quad 3,3$$

2. είναι μεγάλη η διαφορά ανάμεσα στην πραγματική και την επιθυμητή έξοδο του δικτύου:

$$\|\vec{T} - \vec{y}(\vec{I})\| > \epsilon. \quad 3,4$$

Τυπικά η ποσότητα ϵ αντιπροσωπεύει την επιθυμητή ακρίβεια της εξόδου του δικτύου. Σφάλματα μεγαλύτερα από την ποσότητα ϵ σταδιακά διορθώνονται με τη μέθοδο της κατάβασης κλίσης. Η ποσότητα $\delta(t)$ υποδηλώνει τη βεβαιότητα που εμφανίζει το δίκτυο κατά την επανάληψη t των προτύπων εισόδου. Η εκπαίδευση ξεκινά με τη μέγιστη τιμή της ποσότητας $\delta(t) = \delta_{\text{max}}$, τυπικά ίση με το μέγεθος ολόκληρου του διαστήματος μη μηδενικής πυκνότητας πιθανότητας της εισόδου. Η απόσταση $\delta(t)$ συρρικνώνεται μέχρι να φτάσει την ελάχιστη τιμή της δ_{min} . Το δίκτυο θα καταλήξει σε χαρακτηριστικά που είναι μικρότερα της τιμής αυτής.

Η διπλή αυτή συνθήκη είναι απαραίτητη για να κατασκευάσουμε ένα συμπαγές δίκτυο. Εάν εφαρμοστεί μόνο το πρώτο κριτήριο, το δίκτυο θα τοποθετεί συνέχεια κόμβους αντί να χρησιμοποιεί τη μέθοδο κατάβασης κλίσης για να διορθώσει μικρά σφάλματα. Εάν εφαρμοστεί μόνο το δεύτερο κριτήριο της συνθήκης, κόμβοι υψηλής σημασίας θα τοποθετούνται για να αντιπροσωπεύσουν ασήμαντα χαρακτηριστικά, το οποίο είναι δαπανηρό και δε μας συμφέρει. Τοποθετώντας καινούριους κόμβους, το δίκτυο RAN σταδιακά προσεγγίζει όλο και περισσότερο την επιθυμητή συνάρτηση καθώς εκπαιδεύεται. Λιγότεροι κόμβοι χρειάζονται για μια δεδομένη ακρίβεια, αν οι συνάψεις c_{jk} του πρώτου επιπέδου, οι συνάψεις h_{ij} του δεύτερου επιπέδου, και τα κατώφλια γ_i προσαρμόζονται ώστε να μειώσουν το σφάλμα:

$$\mathcal{E} = \|\vec{y} - \vec{T}\|^2. \quad 3,5$$

3.4 Aranfifis- Μοντέλο Προσαρμοζόμενης Δομής με Ανάθεση Πόρων

Το πρωτότυπο αυτό μοντέλο ανάθεσης πόρων που βασίζεται στην αρχιτεκτονική τύπου RAN, και το οποίο θα χρησιμοποιήσουμε ως βάση στη διπλωματική, ονομάζεται ARANFIS (Adaptive Resource Allocating Neural Fuzzy Inference System) [21] και αποτελεί ένα χαρακτηριστικό παράδειγμα συστήματος με προσαρμοζόμενη δομή. Η μέθοδος μη επιβλεπόμενης μάθησης που εφαρμόζεται στη μοντελοποίηση της δομής του προτεινόμενου νευροασαφούς συστήματος, αντίθετα με τον αλγόριθμο RAN που αρχικά δεν περιέχει κρυφούς κόμβους, αρχικοποιεί το σύστημα με ένα μικρό σύνολο κανόνων που περιγράφει χονδρικά το πρόβλημα. Κατά τη διαδικασία της μάθησης, το δίκτυο, βασισμένο σε δυο κριτήρια, είτε αναπτύσσεται προσθέτοντας κρυφούς κόμβους, είτε ρυθμίζει τις ήδη υπάρχουσες παραμέτρους, χρησιμοποιώντας τη μέθοδο της κατάβασης κλίσης των ελάχιστων μέσων τετραγώνων. (least mean square gradient descent). Το πρώτο από τα κριτήρια βασίζεται στο σφάλμα της πρόβλεψης της εξόδου, το οποίο πρέπει να διατηρείται μικρότερο από μια τιμή, και το δεύτερο θέλει τη βεβαιότητα (confidence) του νικητή κανόνα, να είναι μεγαλύτερη από ένα κατώφλι. Αν οι δύο αυτές συνθήκες δεν ικανοποιούνται, τότε ένας νέος κρυφός κόμβος κανόνας προστίθεται στο δίκτυο και οι συνδέσεις του αναπροσαρμόζονται.

Τα βασικά χαρακτηριστικά που διαφοροποιούν το ARANFIS από άλλες μεθόδους είναι τα εξής:

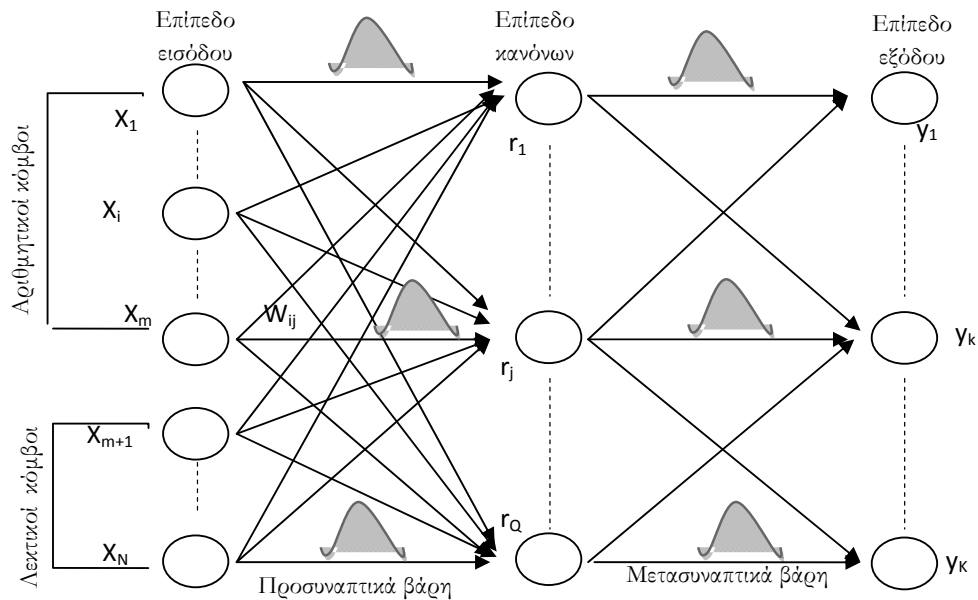
1. Διαθέτει μια αρχιτεκτονική ανάθεσης πόρων, μέσω της οποίας επιδεικνύει ένα είδος προσαρμοστικής συμπεριφοράς, απαραίτητης σε δυναμικά περιβάλλοντα.
2. Μπορεί να χειριστεί τόσο αριθμητικές όσο και γλωσσικές εισόδους, ενώ όλη η πληροφορία που διέρχεται από το δίκτυο είναι ασαφής.
3. Η παράμετρος του ρυθμού μάθησης προσαρμόζεται δυναμικά.
4. Εφαρμόζεται σε ένα ευρύ φάσμα προβλημάτων.

Η αρχιτεκτονική του δικτύου, η οποία είναι παρόμοια με αυτή του SuPFuNIS, παρουσιάζεται στο Σχήμα 3.2. Η προφανής διαφορά ανάμεσα στα δύο συστήματα έχει να κάνει με το γεγονός ότι στο ARANFIS, το πλήθος των εσωτερικών κρυφών κόμβων δεν παραμένει σταθερό, αλλά μεταβάλλεται και διαμορφώνεται δυναμικά κατά τη διάρκεια της εκπαίδευσης.

Η αρχιτεκτονική του ARANFIS έχει τη δυνατότητα να διαχειριστεί τόσο αριθμητικά όσο και γλωσσικά δεδομένα ταυτόχρονα. Τα αριθμητικά δεδομένα ασαφοποιούνται από κόμβους εισόδου που λειτουργούν σαν ασαφοποιητές, ενώ οι συνδέσεις του δικτύου αντιπροσωπεύονται από γκαουσιανές συναρτήσεις συμμετοχής. Η βάση γνώσης των κανόνων εύκολα εισάγεται στο δίκτυο με τη μορφή ασαφών κανόνων EAN- TOTE που ενσωματώνονται στο δίκτυο σαν κρυφοί κόμβοι.

Κατά τη διαδικασία της εκπαίδευσης μπορεί να προστεθούν και άλλοι κρυφοί κόμβοι, εάν οι ήδη υπάρχοντες δεν μπορούν να αντιπροσωπεύσουν τα αριθμητικά δεδομένα.

Από την άλλη, η σωστή αρχικοποίηση του κρυφού επιπέδου είναι απαραίτητη, ώστε να αποφύγουμε τη δημιουργία κόμβων που αντιπροσωπεύουν ασήμαντους κανόνες.



Σχήμα 3.2: Αρχιτεκτονική ARANFIS

3.4.1 Διαδικασία Μάθησης για το ARANFIS

Κατά την εκπαίδευση του ARANFIS χρησιμοποιείται η μέθοδος κατάβασης κλίσης. Το κριτήριο του τετραγωνικού σφάλματος χρησιμοποιείται ως μια παράμετρος επίδοσης της εκπαίδευσης. Το τετραγωνικό σφάλμα σε κάθε επανάληψη υπολογίζεται ως συνήθως από τη σχέση (3.6).

$$e(t) = \frac{1}{2} \sum_{k=1}^p (d_k(t) - y_k(t))^2 \quad 3.6$$

όπου $d_k(t)$ η επιθυμητή έξοδος και $y_k(t)$ η αποσαφαιποιημένη έξοδος στον κόμβο k . Το σφάλμα υπολογίζεται για όλες τις εξόδους p για κάθε πρότυπο εισόδου $x(t)$.

Θεωρούμε πως το δίκτυο, πριν ξεκινήσει η διαδικασία μάθησης, αποτελείται από N κόμβους εισόδου, Q αρχικούς κόμβους κανόνων και K κόμβους εξόδου. Λόγω της δυναμικής δομής του συστήματος κατά τη διαδικασία μάθησης, θέτουμε το πλήθος των κανόνων ίσο με μια μεταβλητή $Q(t)$ που αλλάζει συναρτήσει της επανάληψης t .

Η αποσαφαιποιημένη έξοδος στον κόμβο k θα δίνεται από τη σχέση:

$$y_k(t) = \frac{\prod_{j=1}^{Q(t)} z_j v_{jk}^c v_{jk}^s}{\prod_{j=1}^{Q(t)} z_j v_{jk}^s} \quad 3.7$$

όπου z_j είναι η ισχύς ενεργοποίησης του κανόνα j , K το πλήθος των κόμβων εξόδου και $v_{jk}^c v_{jk}^s$ είναι τα κέντρα και οι διαφορές των μετασυναπτικών βαρών.

Οι ελεύθερες παράμετροι του συστήματος, δηλαδή τα κέντρα και οι διασπορές των προσυναπτικών και μετασυναπτικών βαρών, και οι διασπορές των χαρακτηριστικών εισόδου, μεταβάλλονται με βάση εξισώσεις ανανέωσης της παρακάτω μορφής:

$$w_{ij}^c(t+1) = w_{ij}^c(t) - \eta(t) \cdot a_j \frac{\partial e(t)}{\partial w_{ij}^c(t)} \quad 3.8$$

$$v_{jk}^c(t+1) = v_{jk}^c(t) - \eta(t) \cdot a_j \frac{\partial e(t)}{\partial v_{jk}^c(t)} \quad 3.9$$

$$w_{ij}^\sigma(t+1) = w_{ij}^\sigma(t) - \eta(t) \cdot a_j \frac{\partial e(t)}{\partial w_{ij}^\sigma(t)} \quad 3.10$$

$$v_{jk}^\sigma(t+1) = v_{jk}^\sigma(t) - \eta(t) \cdot a_j \frac{\partial e(t)}{\partial v_{jk}^\sigma(t)} \quad 3.11$$

$$x_i^\sigma(t+1) = x_i^\sigma(t) - \eta(t) \cdot a_j \frac{\partial e(t)}{\partial x_i^\sigma(t)} \quad 3.12$$

όπου $\eta(t)$ είναι ο προσαρμοζόμενος ρυθμός μάθησης, άλλο ένα χαρακτηριστικό του ARANFIS που το κάνει να διαφέρει από τα υπόλοιπα συστήματα γινομένου και να εμφανίζει μεγαλύτερη προσαρμοστικότητα. Ο υπολογισμός των μερικών παραγώγων που εμφανίζονται στις παραπάνω εξισώσεις δίνεται αναλυτικά στο Παράρτημα Α στο τέλος της διπλωματικής.

3.4.2 Προσαρμογή ρυθμού μάθησης:

Η επιλογή της τιμής του ρυθμού μάθησης, η , αποτελεί καθοριστικό παράγοντα στην απόδοση του δικτύου, αφού είναι στενά συνδεδεμένος με την ταχύτητα σύγκλισης. Ο κατάλληλος χειρισμός του learning rate και η προσαρμογή του κατά τη διάρκεια της εκπαίδευσης μπορεί να οδηγήσει σε καλύτερα αποτελέσματα, όπως αποδεικνύεται από σχετικές μελέτες στη βιβλιογραφία.[17, 23]. Στη διαδικασία μάθησης του ARANFIS ο ρυθμός μάθησης υπολογίζεται σύμφωνα με την υπόθεση ότι το σύνολο των δεδομένων μπορεί να χωριστεί σε υποσύνολα με παρόμοια πρότυπα. Εφόσον το προτεινόμενο δίκτυο έχει ήδη διαμορφώσει κάποιες ομάδες προτύπων για την αρχικοποίησή του, ο ρυθμός μάθησης μπορεί να οριστεί ως:

$$\eta(t) = \frac{\beta}{\sqrt{N_1^2(t) + N_2^2(t) + \dots + N_{q(t)}^2(t) + 1}} \quad 3.13$$

όπου β μια εμπειρική σταθερά και $N_j(t)$ το πλήθος των δεδομένων εισόδου, μέχρι την επανάληψη t , για τα οποία ο κρυφός κόμβος j ήταν ο κόμβος νικητής.

Διαδικασία εισαγωγής νέου κόμβου

Η διαδικασία εισαγωγής νέου κόμβου, θεωρείται σχετικά απλή. Θεωρούμε το ζεύγος δεδομένων εκπαίδευσης $x(t)$, $y(t)$. Όταν το πρότυπο εισόδου $x(t)$ δεν ενεργοποιεί ικανοποιητικά κανένα από τους υπάρχοντες κανόνες και το σφάλμα πρόβλεψης είναι σημαντικό, ένας νέος κανόνας δημιουργείται και προστίθεται στο σύστημα, όπου πλέον ισχύει $Q(t) = Q(t-1) + 1$.

Συγκεκριμένα πρέπει να ισχύει:

$$\max_j \{z_j\} < \delta \quad 3.14$$

Και

$$|d_k(t) - y_k(t)| > \varepsilon \quad 3.15$$

3.5 Αξιολόγηση και Δυναμική επιλογή κανόνων- Προτεινόμενη Μέθοδος

Μέθοδος Αξιολόγησης και Δυναμικής Επιλογής Κανόνων

Η εύρεση του βέλτιστου πλήθους κανόνων που απαιτούνται για την επίλυση ενός προβλήματος είναι ένα ανοιχτό ερευνητικό θέμα στο χώρο των νευροασαφών συστημάτων. Μια διαφορετική προσέγγιση είναι αυτή που προτείνει την αξιολόγηση των κανόνων, ανεξάρτητα από το αρχικό πλήθος τους, με στόχο την αφαίρεση ή/και τον επαναπροσδιορισμό κάποιων από αυτούς, ώστε να απλοποιηθεί η βάση των κανόνων και να αντιμετωπιστεί εν μέρει η «κατάρα της διάστασης». Διάφοροι τέτοιοι αλγόριθμοι έχουν προταθεί στη βιβλιογραφία, οι οποίοι προσπαθούν να εκμεταλλευτούν τα θετικά στοιχεία, των γενετικών αλγορίθμων, της ενισχυτικής μάθησης, ή άλλων τεχνικών υπολογιστικής ευφυΐας [11, 16, 19]. Γεγονός όμως είναι ότι όλες αυτές οι ιδέες έχουν ως βασικό στόχο την εξάλειψη των πιο αδύναμων κόμβων από το δίκτυο (είτε πρόκειται για κόμβους κανόνων ή εισόδου). Αυτό συνεπάγεται τη μείωση του υπολογιστικού κόστους, αλλά και την απώλεια κάποιων στοιχείων πληροφορίας. Η απουσία αυτής της πληροφορίας μπορεί να έχει ελάχιστη αρνητική επίδραση στη συνολική απόδοση του συστήματος, ενδεχομένως όμως σε κάποιες περιπτώσεις η έλλειψη αυτή να αποδειχθεί κρίσιμη.

Αξιολόγηση και δυναμική επιλογή κανόνων με μέτρο την ευαισθησία

Η κεντρική ιδέα για τη δυναμική επιλογή κανόνων που συνθέτουν το αποτέλεσμα σε ένα νευροασαφές δίκτυο γινομένου, βασίζεται σε μια ιεραρχική δομή του συνόλου των ασαφών κανόνων. Σύμφωνα με τη δομή αυτή η υψηλότερη βαθμίδα με τους λιγότερους κανόνες ενεργοποιείται πρώτη και, αν δεν προκύψει ένα επιθυμητό αποτέλεσμα, χρησιμοποιείται και η δεύτερη βαθμίδα με τους παραπάνω κανόνες. Αν και αυτή δε δώσει ένα επιθυμητό αποτέλεσμα, ενεργοποιείται και η τρίτη βαθμίδα με τους παραπάνω κανόνες κ.ό.κ., μέχρι να ενεργοποιηθεί και η τελευταία βαθμίδα που περιέχει το σύνολο της πληροφορίας. Σκοπός αυτής της μεθοδολογίας είναι η εξοικονόμηση πόρων και η μείωση του κόστους κατά τη

διάρκεια λειτουργίας του συστήματος. Με άλλα λόγια το εν λόγω μοντέλο κάνει χρήση μόνο εκείνων των κόμβων του επιπέδου κανόνων που χρειάζεται κάθε φορά χωρίς να αποκλείει εντελώς κάποιους κανόνες από την επίλυση του προβλήματος.

Η μεθοδολογία που προτείνεται εκμεταλλεύεται τα πλεονεκτήματα της ανάλυσης με βάση την ευαισθησία και χρησιμοποιεί ένα μέτρο ευαισθησίας για να εξάγει το βαθμό σπουδαιότητας του κάθε κανόνα ως προς την έξοδο του συστήματος. Στη συνέχεια δημιουργούνται υποσύνολα κανόνων, με βάση το κριτήριο διαχωρισμού και καταλήγουμε σε μια ιεραρχική δομή, όπου κάθε επόμενο επίπεδο περιλαμβάνει και τους κανόνες του προηγούμενου επιπέδου. Η επιλογή του επιπέδου προς χρήση πραγματοποιείται με τη βοήθεια ενός μέτρου εμπιστοσύνης.

Να σημειωθεί ότι όλη η διαδικασία αξιολόγησης και ιεράρχησης των κανόνων εφαρμόζεται αφού έχει ολοκληρωθεί η διαδικασία εκπαίδευσης και έχουν διαμορφωθεί οι παράμετροι του συστήματος.

3.6 Αξιολόγηση και Δυναμική Επιλογή Κανόνων στο μοντέλο ARANFIS

Όπως είδαμε το μοντέλο προσαρμοζόμενης δομής με ανάθεση πόρων (Aranfis) εμφανίζει πολλά ιδιαίτερα χαρακτηριστικά τα οποία εξυπηρετούν σε μεγάλο βαθμό τις ανάγκες αυτής της διπλωματικής εργασίας. Η ικανότητα επεξεργασίας ασαφούς πληροφορίας, τόσο σε μορφή αριθμητικών όσο και γλωσσικών μεταβλητών, σε συνδυασμό με τη δυνατότητα που προσφέρει το ARANFIS για δυναμική διαχείριση δομής, το κάνουν το ιδανικό μοντέλο νευροασαφούς συστήματος για τους στόχους που θέσαμε. Μπορούμε λοιπόν να εκμεταλλευτούμε όλες αυτές τις δυνατότητες που μας προσφέρει και να εφαρμόσουμε και νέες μεθόδους και τεχνικές, ώστε να πετύχουμε ένα νευροασαφές σύστημα που θα μπορεί να χρησιμοποιηθεί σε ένα εύρος εφαρμογών της καθημερινότητας και θα ικανοποιεί και τις ανάγκες μας για σχεδιαστική οικονομία, χαμηλό σχεδιαστικό, ενεργειακό και λειτουργικό κόστος, και σωστή διαχείριση και κατανάλωση των διαθέσιμων πόρων.

Έχουν γίνει αρκετές προσπάθειες και πολλές τεχνικές έχουν προταθεί για τη δημιουργία ενός δικτύου με το κατάλληλο μέγεθος και τη βέλτιστη δομή, όχι μόνο για νευροασαφή συστήματα, αλλά και για τις υπόλοιπες κατηγορίες νευρωνικών δικτύων. Καμιά πρακτική, βέβαια, μέθοδος δεν υπάρχει για να καθορίσει *a priori* το κατάλληλο μέγεθος και τις συνδέσεις για ένα δίκτυο [14]. Η πιο σωστή και πολλά υποσχόμενη προσέγγιση είναι, με αφετηρία ένα μεγάλο, πλήρως συνδεδεμένο δίκτυο, και με χρήση διαφόρων μεθόδων αποκοπής και τεχνικών επεξεργασίας (κανονικοποίηση, προετοιμασία δεδομένων και κόμβων), να μειωθεί η διασπορά του δικτύου και να βελτιωθεί ως προς τη γενίκευση.

Ένας από τους πολλούς αλγόριθμους που υπάρχουν για την αποκοπή των γενικών παραμέτρων και τη βελτίωση της γενίκευσης του δικτύου μετά από επιβλεπόμενη μάθηση, είναι των Mozer και Smolesky (1989) [14], ο οποίος απομακρύνει τους νευρώνες που έχουν τη μικρότερη επίδραση στο σφάλμα εξόδου. Είναι όμως δαπανηρός και δε λαμβάνει υπόψη του τους συσχετισμούς ανάμεσα στους νευρώνες. Η εξάλειψη των μικρών βαρών δεν υπολογίζεται απαραίτητα σαν επίδραση ενός βάρους στο σφάλμα εξόδου. Ένας άλλος αλγόριθμος, ο OBD (Optimal Brain Damage, Le Cun et al.1990) [14], αφαιρεί τα βάρη εκείνα που επηρεάζουν λιγότερο το σφάλμα εκπαίδευσης, βασιζόμενος σε μια διαγωνιοποίηση του πίνακα Hessian. Πρόκειται όμως για μια εσφαλμένη προσέγγιση, η οποία μπορεί να οδηγήσει στην αφαίρεση των λάθος βαρών. Η μέθοδος αυτή απαιτεί, επιπλέον, επανεκπαίδευση του καινούριου και ελαττωμένου δικτύου, το οποίο υπολογιστικά έχει υψηλό κόστος [14].

Στην παρούσα διπλωματική εργασία έγινε προσπάθεια συνδυασμού της ανάλυσης ευαισθησίας, εφαρμοσμένης όμως με έναν καινούριο τρόπο στο δίκτυο, με το μοντέλο του ARANFIS, και επιπλέον προσπάθεια ελέγχου του αριθμού των κρυφών κόμβων κανόνων, τόσο κατά τη διάρκεια της εκπαίδευσης του δικτύου, όσο και μετά, κατά την εφαρμογή του. Όσον αφορά στην ανάλυση ευαισθησίας, όπως αναφέρθηκε και προηγουμένως, την εφαρμόσαμε όχι στο επίπεδο της εισόδου σε σχέση με την έξοδο, αλλά στο επίπεδο των κανόνων. Θεωρήσαμε την ισχύ ενεργοποίησης κάθε κανόνα ως μεταβλητή προς μελέτη και βρήκαμε την επίδρασή της στην έξοδο του δικτύου. Με τον τρόπο αυτό, εισχωρήσαμε στην εσωτερική δομή του δικτύου και προσπαθήσαμε να τη διαχειριστούμε και να τη μεταβάλλουμε, δηλαδή δεν περιοριστήκαμε στη μελέτη της επίδρασης της εισόδου ως προς την έξοδο, μένοντας στις εξωτερικές παραμέτρους.

Η νέα αυτή πρόταση δεν επέφερε κάποια επιβάρυνση ούτε σε υπολογιστικό, ούτε σε σχεδιαστικό κόστος, αφού απαιτείται μόνο ο υπολογισμός κάποιων μερικών παραγώγων, όπως και στην κλασική εφαρμογή της ανάλυσης ευαισθησίας. Επομένως, ο στόχος για «οικονομικό» σχεδιασμό και λειτουργία του δικτύου δεν εμποδίζεται. Ο συσχετισμός αυτός ανάμεσα στην ισχύ ενεργοποίησης κάθε κανόνα και της εξόδου, γίνεται αμέσως μετά το τέλος των εποχών εκπαίδευσης. Αφού το δίκτυο έχει εκπαιδευτεί, υπολογίζονται οι μερικές παράγωγοι της ισχύος ενεργοποίησης κάθε κανόνα ως προς την έξοδο και προκύπτει ποιοι κανόνες είναι σημαντικοί για την έξοδο του δικτύου.

Με βάση το αποτέλεσμα που θα προκύψει από την ανάλυση ευαισθησίας, προχωράμε και στην αξιολόγηση και δυναμική επιλογή των κανόνων.

Το ARANFIS, ως σύστημα δυναμικής διαχείρισης πόρων, έχει τη δυνατότητα να προσθέσει νέους κανόνες, έτσι ώστε να ικανοποιηθούν τα απαραίτητα κριτήρια για την απόδοση του δικτύου. Στόχος είναι, από τους κανόνες αυτούς που προσθέτει το σύστημά μας, να μπορέσουμε να διακρίνουμε ποιοί είναι απαραίτητοι ή ιδιαίτερα σημαντικοί για την έξοδο του δικτύου, ώστε να τους χρησιμοποιήσουμε. Οι υπόλοιποι κανόνες που φαίνεται να είναι μικρότερης σημασίας, και άρα δε θα έχουν μεγάλη επίδραση στην έξοδο του δικτύου,

μπορούν να μη χρησιμοποιηθούν. Δεν είναι απαραίτητη η αποκοπή τους, αλλά παραμένουν στο σύστημα ανενεργοί και δε χρησιμοποιούνται, τουλάχιστον όχι για όλες τις εξόδους. Έτσι, από τη στιγμή που δεν αποκόπτεται κάτι από την εσωτερική δομή του δικτύου, δεν υπάρχει και απώλεια πληροφορίας.

Μεγάλη σημασία έχει και το κριτήριο με το οποίο θα αποφασίσει το δίκτυο αν κάποιος από τους κανόνες είναι περισσότερο ή λιγότερο σημαντικός.

Η μέθοδος που προτείνουμε στηρίζεται στο γεγονός πως ο ορισμός της ευαισθησίας για την περίπτωση μας, αφορά στο ρυθμό μεταβολής της εξόδου ως προς την ισχύ ενεργοποίησης.

Αναλυτικότερα:

Έστω ένα δίκτυο με N κόμβους εισόδου, Q κόμβους κανόνων και K κόμβους εξόδου. Ορίζουμε την *ευαισθησία κανόνα (Rule Sensitivity)*, για ένα εκπαιδευμένο ζευγάρι p κανόνα-εξόδου, ως εξής:

$$RS_{kj}^{(p)} = \frac{\partial y_k}{\partial z_j} \quad 3.16$$

όπου $k= 1..K$, $j= 1..Q$. Επειδή κάθε ζεύγος p παράγει έναν διαφορετικό Ιακωβιανό πίνακα, υπολογίζουμε το μέτρο RS για όλο το σύνολο δεδομένων και λαμβάνουμε τον πίνακα μέσης τιμής:

$$RS_{kj,avg} = \sqrt{\frac{\sum_{p=1}^P [RS_{kj}^{(p)}]^2}{N}} \quad 3.17$$

όπου P το πλήθος των προτύπων που λαμβάνουν μέρος στην εκπαίδευση.

Για να έχουμε σωστή σύγκριση μεταξύ των μεταβλητών, είναι απαραίτητο να μετασχηματίσουμε τις τιμές ενεργοποίησης και εξόδου στην ίδια κλίμακα, χρησιμοποιώντας τον τύπο:

$$\overline{RS_{kj,avg}} = RS_{kj,avg} \frac{(\max_{p=1..P}\{z_j^{(p)}\} - \min_{p=1..P}\{z_j^{(p)}\})}{(\max_{p=1..P}\{y_k^{(p)}\} - \min_{p=1..P}\{y_k^{(p)}\})} \quad 3.18$$

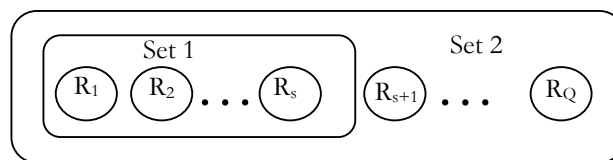
Επομένως ο τελικός πίνακας $RS_{kj,avg}$ περιγράφει τη σχέση του κάθε κανόνα ως προς τη μεταβλητή εξόδου. Το επιθυμητό, όμως είναι να προσδιορίσουμε τη σημαντικότητα της ισχύος ενεργοποίησης κάθε κανόνα ως προς τη γενική έξοδο του συστήματος, και για το λόγο αυτό ορίζουμε το *βαθμό σπουδαιότητας* Φ_j του κανόνα j ως προς την απόκριση του συστήματος ως:

$$\Phi_j = \max_{k=1..K}\{RS_{kj,avg}\} \quad 3.19$$

Το διάνυσμα σπουδαιότητας Φ , δηλαδή οι βαθμοί σπουδαιότητας του συνόλου των κανόνων, ταξινομούνται με σειρά απόδοσης, έτσι ώστε ο πιο σημαντικός κανόνας, q_1 να είναι ο πρώτος και ο λιγότερο σημαντικός q_Q ο τελευταίος.

Με βάση τα μεγέθη που προκύπτουν από τις παραπάνω εξισώσεις, βγάζουμε συμπεράσματα για το ποιοι κανόνες μπορούν να θεωρηθούν σημαντικοί για το σύστημά μας, οπότε και να χρησιμοποιηθούν οπωσδήποτε για τη λειτουργία του συστήματος, και ποιοι κανόνες μπορούν να θεωρηθούν δευτερεύουσας σημασίας. Οι κανόνες, που δε θεωρούνται τόσο σημαντικοί, μπορούν να μη χρησιμοποιηθούν, τουλάχιστον αρχικά, αν το σύστημα εμφανίζει μια καλή, σχετικά, απόδοση και χωρίς αυτούς. Στόχος μας είναι, δηλαδή, να κάνουμε μια διαίρεση των κανόνων σε ιεραρχημένα υποσύνολα, έτσι ώστε να μπορούμε να τους διαχειριστούμε δυναμικά κατά τη λειτουργία του συστήματος.

Η μέθοδος που προτείνεται, είναι να ιεραρχήσουμε τους κανόνες που θα προκύψουν τελικά στο σύστημά μας, μετά και την προσθήκη κρυφών κόμβων κατά τη διάρκεια της εκπαίδευσης, με βάση την ισχύ ενεργοποίησής τους. Και πιο συγκεκριμένα, να εκμεταλλευτούμε και την ανάλυση ευαισθησίας όπως την εφαρμόσαμε, και να χρησιμοποιήσουμε ως κριτήριο σημαντικότητας των κανόνων, το ρυθμό μεταβολής της εξόδου προς την ισχύ ενεργοποίησης z_j , το μέγεθος, δηλαδή, Φ_j , που υπολογίστηκε στη σχέση 3.19. Αν λοιπόν το μέτρο Φ_j είναι ίσο με τη μονάδα, τότε η μεταβολή της εξόδου είναι σε αναλογία ένα προς ένα με οποιαδήποτε μεταβολή της ισχύος ενεργοποίησης z_j του κανόνα j . Είναι τότε λογικό επακόλουθο να θεωρήσουμε πως, αν η τιμή Φ_j είναι μεγαλύτερη της μονάδας, αυτό σημαίνει πως ο αντίστοιχος κανόνας είναι πιο σημαντικός για την έξοδο, από ότι στην περίπτωση που η τιμή είναι μικρότερη της μονάδας. Με βάση αυτό το κριτήριο χωρίζουμε το σύνολο των κανόνων σε δύο υποσύνολα: την κύρια βάση κανόνων και τη δευτερεύουσα, όπως φαίνεται και στο σχήμα 3.3:



Σχήμα 3.3: Ιεραρχική δομή κανόνων- Χωρισμός σε δυο υποσύνολα

Η κύρια βάση κανόνων θα χρησιμοποιηθεί έτσι κι αλλιώς από το σύστημα, ενώ η δευτερεύουσα θα προστεθεί μόνο αν το σύστημα δεν μπορεί να ικανοποιήσει τα επιβαλλόμενα κριτήρια.

3.7 Περιγραφή Λειτουργίας του μοντέλου ARANFIS με την προτεινόμενη μεθοδολογία

Η λειτουργία του συστήματος με βάση τη μεθοδολογία που αναλύεται στην ενότητα 3.6, στηρίζεται στο αν η βεβαιότητα της εξόδου θα ικανοποιήσει ή όχι ένα κριτήριο εμπιστοσύνης (confidence measure). Το προτεινόμενο μέτρο εμπιστοσύνης C για τα προβλήματα ταξινόμησης είναι μια επέκταση της μεθόδου που περιγράφεται στο [5], όπου ως κριτήριο αξιοπιστίας δίνεται η διαφορά μεταξύ των δυο μεγαλύτερων τιμών εξόδου. Στην παρούσα διπλωματική, η διαφορά αυτή πολλαπλασιάζεται με τη μεγαλύτερη τιμή, ώστε να περιλαμβάνει και το μέγεθος της ισχύος της εξόδου:

$$C = (Y_{win1} - Y_{win2}) * Y_{win1} \quad 3.20$$

όπου Y_{win1} είναι η αριθμητική τιμή του κόμβου εξόδου με τη μεγαλύτερη ισχύ, και Y_{win2} η αριθμητική τιμή του κόμβου με τη δεύτερη μεγαλύτερη τιμή. Αν το μέτρο εμπιστοσύνης είναι μικρότερο από ένα συγκεκριμένο κατώφλι εμπιστοσύνης που θέτουμε εμείς, το σύστημα θεωρεί ότι το κύριο σύνολο κανόνων δεν επαρκεί και χρησιμοποιεί και το δευτερεύον σύνολο κανόνων, ουσιαστικά δηλαδή όλους τους κανόνες που προέκυψαν από το στάδιο της εκπαίδευσης. Πρακτικά, το κατώφλι εμπιστοσύνης που θέτουμε για το σύστημα, καθορίζει το ποσοστό στο οποίο χρησιμοποιείται η κύρια βάση κανόνων του συστήματος και ταυτόχρονα προσφέρει μια ευελιξία στη διαμόρφωση της τρέχουσας δομής του δικτύου, αφού από αυτό εξαρτάται το πλήθος των κρυφών κόμβων κανόνων που θα χρησιμοποιηθούν.

Κεφάλαιο 4: Πειραματική Μελέτη

Η πειραματική μελέτη περιλαμβάνει την εκτέλεση δοκιμών σε τέσσερα διαφορετικά σύνολα δεδομένων και συγκεκριμένα στα Breast Cancer, Ionosphere, Pima Indians και Vehicle (Παράρτημα Β). Για κάθε σύνολο δεδομένων πραγματοποιήσαμε πειράματα αλλάζοντας τιμές σε σημαντικές παραμέτρους του συστήματος και παρατηρώντας κάθε φορά την αντίδραση του δικτύου σε αυτές τις μεταβολές. Με βάση αυτές τις δοκιμές καταλήξαμε σε κάποια συμπεράσματα για την αποτελεσματικότητα της μεθοδολογίας που εφαρμόσαμε, αλλά και για την επίδραση των παραμέτρων στο σύστημά μας. Για κάθε σύνολο δεδομένων εκτελέστηκαν δέκα επαναλήψεις για κάθε περίπτωση, έτσι ώστε να μπορούμε να συγκρίνουμε τα αποτελέσματα με βάση τους αντίστοιχους μέσους όρους.

Για κάθε σύνολο δεδομένων, εκπαιδεύσαμε το δίκτυο για διάστημα 100 εποχών, με ρυθμό μάθησης 0,001. Η αρχικοποίηση των βαρών έγινε με τυχαίο τρόπο, ενώ ο αρχικός αριθμός των κρυφών κόμβων κανόνων για κάθε σύνολο δεδομένων ορίστηκε τουλάχιστον ίσος με το πλήθος των κατηγοριών στις οποίες χωρίζονται οι έξοδοι για κάθε σύνολο δεδομένων: συγκεκριμένα δύο αρχικοί κανόνες για τα σύνολα δεδομένων Breast Cancer, Ionosphere, και Pima Indians, και τέσσερις αρχικοί κανόνες για το σύνολο δεδομένων Vehicle. Το κατώφλι εμπιστοσύνης καθορίστηκε για τις περισσότερες περιπτώσεις ίσο με 0,5, όμως αυτήν την τιμή τη μεταβάλαμε πολλές φορές κατά τη διάρκεια των δοκιμών, για να παρατηρήσουμε τα αντίστοιχα αποτελέσματα.

Μετά την εκπαίδευση του δικτύου, προχωρήσαμε στην αξιολόγηση (testing). Για κάθε σύνολο δεδομένων πραγματοποιήσαμε ένα απλό πέρασμα του συνόλου testing των δεδομένων από το εκπαιδευμένο δίκτυο, όπου παρατηρήσαμε την απόδοση και το χρόνο εκτέλεσης του δικτύου, ώστε να συμπεράνουμε κατά πόσο έγινε σωστά η εκπαίδευσή του.

Στη συνέχεια πραγματοποιήσαμε και μια ακόμα προσομοίωση (simulation) του δικτύου, χρησιμοποιώντας όμως αυτήν τη φορά τη μέθοδο αξιολόγησης και ιεράρχησης των κανόνων σε δυο υποσύνολα, (κεφάλαιο 3), με βάση το βαθμό σπουδαιότητας του κάθε κανόνα, μέθοδο που χωρίζει τους κανόνες σε ένα κύριο υποσύνολο κανόνων και ένα δευτερεύον. Την απόδοση και το χρόνο εκτέλεσης αυτής της προσομοίωσης, τα συγκρίναμε με τα αντίστοιχα μεγέθη της απλής προσομοίωσης, ώστε να καταλήξουμε σε ένα συμπέρασμα σχετικά με τη χρησιμότητα της προτεινόμενης μεθόδου. Πέρα από το χρόνο και την απόδοση του δικτύου, στην προσομοίωση με την προτεινόμενη μέθοδο, μπορούμε να παρατηρήσουμε και άλλα μεγέθη, όπως σε τι ποσοστό χρησιμοποιήθηκε μόνο το κύριο υποσύνολο κανόνων και σε ποιο χρησιμοποιήθηκαν όλοι οι κανόνες που πρόσθεσε το δίκτυο κατά την εκπαίδευση. Από αυτά τα μεγέθη μπορούμε να βγάλουμε πολύ χρήσιμα συμπεράσματα σχετικά με το τι προσφέρει η μέθοδος που προτείνουμε στη μείωση του σχεδιαστικού και υπολογιστικού κόστους του συστήματός μας, αλλά και για τη μορφή του συνόλου δεδομένων, και άρα τις εφαρμογές στις οποίες μπορεί να χρησιμοποιηθεί ο συνδυασμός που προτείνουμε (μοντέλο δυναμικής διαχείρισης δομής και ανάθεσης πόρων- ARANFIS- και ιεράρχηση κανόνων με βάση την ευαισθησία).

Κατά την εκτέλεση των δοκιμών, η ανάλυση ευαισθησίας εφαρμόστηκε στο σύστημα μετά το τέλος της εκπαίδευσης, αφού δηλαδή είχε εκπαιδευτεί το δίκτυο, και είχαν προστεθεί όλοι οι κανόνες που χρειαζόταν το σύστημα με βάση τα επιβαλλόμενα κριτήρια. Επιπλέον, κατά την προσομοίωση του δικτύου, δε γίνεται εκ νέου υπολογισμός της ευαισθησίας, αφού κάτι τέτοιο θα ήταν απλά χρονοβόρο, χωρίς να προσφέρει κάτι ουσιαστικό στη μελέτη μας και στη λειτουργία του συστήματος.

Στους πίνακες 4.1 έως 4.12, φαίνονται τα αποτελέσματα των πειραματικών δοκιμών για κάθε σύνολο δεδομένων.

Σύνολο Δεδομένων	Αρχικός αριθμός κανόνων	ϵ (M.O)	δ (M.O)	Τελικός αριθμός κανόνων (M.O)	Απόδοση Εκπαίδευσης (%,M.O)
Breast Cancer	2	0,36	0,5	14,5	96,75
Ionosphere	2	0,36	0,5	80,3	88,3
Pima Indians	2	0,36	0,5	5,5	76,8

Vehicle	4	0,47	0,43	68,46	46,18

Πίνακας 4.1: Αποτελέσματα πειραμάτων για την εκπαίδευση του συστήματος

Ο πίνακας 4.1 δείχνει τα αποτελέσματα από την εκπαίδευση για κάθε ένα από τα τέσσερα σύνολα δεδομένων που χρησιμοποιήσαμε. Ο αρχικός αριθμός κανόνων είναι οι κανόνες που βάζουμε εμείς αρχικά στο σύστημα, πριν αρχίσει η εκπαίδευση. Το ϵ είναι μία από τις δύο συνθήκες που περιγράφηκαν στο κεφάλαιο 3 για την προσθήκη νέου κόμβου στο δίκτυο, η οποία δίνεται από τη σχέση 3.15. Η τιμή που φαίνεται στον πίνακα είναι ο μέσος όρος από τα δέκα πειράματα που εκτελέσαμε για κάθε σύνολο δεδομένων. Το σύστημα συγκρίνει σε κάθε επανάληψη την προκύπτουσα τιμή του ϵ , με τη δοθείσα (το μέσο όρο που φαίνεται στον πίνακα 4.1) και ανάλογα προσθέτει ή όχι κανόνα. Η διαδικασία αυτή πραγματοποιείται αφού ελεγχθεί και η δεύτερη συνθήκη, το δ , που φαίνεται στην επόμενη στήλη του πίνακα 4.1. Και σε αυτήν την περίπτωση, το σύστημα συγκρίνει την τιμή που προκύπτει κάθε φορά για το δ από το σύστημα με αυτήν που έχουμε θέσει ως κατώφλι στη δοκιμή μας. Με βάση αυτές τις συνθήκες προκύπτει και ο τελικός αριθμός κανόνων για το σύστημα. Όπως παρατηρούμε, στα σύνολα δεδομένων breast cancer και pima Indians, το σύστημα δεν προσθέτει μεγάλο αριθμό κανόνων (12,5 και 3,5 κανόνες αντίστοιχα κατά μέσο όρο). Αντίθετα, στα σύνολα ionosphere και vehicle, το σύστημα προσθέτει κατά μέσο όρο 78,3 και 64,46 κανόνες, αυξάνοντας κατά πολύ τον τελικό αριθμό. Η διαφορά αυτή είναι αποτέλεσμα των διαφορετικών χαρακτηριστικών των συνόλων δεδομένων, αφού τα κριτήρια προσθήκης νέων κανόνων (ϵ , δ) δε διαφέρουν αισθητά. Έτσι λοιπόν φαίνεται ότι τα δυο αυτά σύνολα δεδομένων, είναι πιο πολύπλοκα και «μπερδεμένα», οπότε το σύστημα προσθέτει περισσότερους κανόνες κατά την εκπαίδευση προκειμένου να τα ταξινομήσει.

Στη συνέχεια παρουσιάζονται τα αποτελέσματα που προκύπτουν από την προσομοίωση του συστήματος, αρχικά με απλό πέρασμα των δεδομένων, και έπειτα με εφαρμογή της μεθόδου δυναμικής διαχείρισης και ιεράρχησης κανόνων.

Για την απλή προσομοίωση του εκπαιδευμένου δικτύου λαμβάνουμε τα αποτελέσματα του πίνακα 4.2:

Σύνολο Δεδομένων	Απόδοση προσομοίωσης	Χρόνος εκτέλεσης
Breast Cancer	98,23	4,24
Ionosphere	87,23	37
Pima Indians	80,1	1,2

Vehicle	39,93	27,9

Πίνακας 4.2: Αποτελέσματα απλής προσομοίωσης εκπαιδευμένου συστήματος

Για την προσομοίωση του εκπαιδευμένου δικτύου με εφαρμογή της μεθόδου ιεράρχησης και διαχείρισης κανόνων, τα αποτελέσματα καταγράφονται στον πίνακα 4.3:

Σύνολο Δεδομένων	Πλήθος κανόνων με $\Phi_i > 1$	Χρόνος εκτέλεσης	Κατώφλι εμπιστοσύνης	Απόδοση Συστήματος	Κύριο υποσύνολο κανόνων (%)	Και δευτερεύον υποσύνολο (%)
Breast Cancer	4,1	2,28	0,7	95,92	53,75	46,32
Ionosphere	14,3	22,95	0,6	85,8	61,93	38,1
Pima Indians	2,4	0,7243	0,5	72,67	68,6	30,7
Vehicle	61,15	51,69	0,5	37,27	12,04	87,96

Πίνακας 4.3: Αποτελέσματα προσομοίωσης εκπαιδευμένου συστήματος με ιεράρχηση κανόνων με κριτήριο την ευαισθησία.

Συγκεντρωτικά, οι αποδόσεις των δύο δικτύων και οι χρόνοι εκτέλεσης φαίνονται στον πίνακα 4.4 για να διευκολυνθεί η διαδικασία της σύγκρισης των δυο συστημάτων:

Σύνολο Δεδομένων	Απόδοση Συστήματος Χωρίς Δυναμική Αξιολόγηση Κανόνων (% , M.O)	Απόδοση Συστήματος με Μέθοδο Δυναμικής Αξιολόγησης Κανόνων (% , M.O)	Χρόνος εκτέλεσης1- Χωρίς Δυναμική Αξιολόγηση Κανόνων (sec, M.O)	Χρόνος εκτέλεσης2- Με Δυναμική Αξιολόγηση Κανόνων (sec, M.O)
Breast Cancer	98,23	95,92	4,24	2,28
Ionosphere	87,23	85,8	37	22,95

Pima Indians	80,1	72,67	1,2	0,7243
Vehicle	39,93	37,27	27,9	51,69

Πίνακας 4.4: Συγκριτικός πίνακας για τις δυο προσομοιώσεις: Το σύστημα χωρίς δυναμική αξιολόγηση και ιεράρχηση κανόνων συγκρίνεται με το σύστημα όπου εφαρμόστηκε η προτεινόμενη μέθοδος δυναμικής αξιολόγησης κανόνων και χωρισμού τους σε δυο υποσύνολα.

Όπως παρατηρούμε από τον πίνακα 4.4, η απόδοση του δικτύου με την εφαρμογή της δυναμικής διαχείρισης των κανόνων, είναι ελάχιστα μικρότερη, κατά μέσο όρο, γεγονός όχι θετικό. Όμως αν αναλογιστούμε την ευελιξία που μας προσφέρει το σύστημα με τη δυναμική διαχείριση δομής και την εξοικονόμηση σχεδιαστικού και λειτουργικού κόστους, καταλήγουμε στο συμπέρασμα ότι, συνολικά, το σύστημα με τη δυναμική διαχείριση δομής είναι πολύ πιο αποδοτικό. Επιπλέον, αν παρατηρήσουμε το χρόνο εκτέλεσης για κάθε ένα από τα δυο συστήματα, διαπιστώνουμε ότι, με την ιεράρχηση των κανόνων και το χωρισμό τους σε δυο υποσύνολα, από τα οποία μπορεί να χρησιμοποιηθεί και μόνο το κύριο, εξοικονομούμε αρκετό χρόνο. Ακόμα και αν μιλάμε για δευτερόλεπτα ή δέκατα του δευτερολέπτου, η γρηγορότερη εκτέλεση είναι ένα πολύ θετικό στοιχείο για ένα σύστημα, αφού με την εξοικονόμηση χρόνου, εξοικονομείται υπολογιστικό, και κυρίως, ενεργειακό κόστος. Καλύπτουμε έτσι μια ανάγκη που παρουσιάζεται πολύ έντονη τα τελευταία χρόνια, αυτή για την εξοικονόμηση ενέργειας (και κατά συνέπεια περιβαλλοντικής προστασίας).

Η παραπάνω παρατήρηση δεν προκύπτει από τη μελέτη του συνόλου δεδομένων Vehicle, όπου η απόδοση εμφανίζεται σε όλες τις περιπτώσεις αρκετά μειωμένη, ακόμα και αν συγκριτικά το σύστημα με τη δυναμική διαχείριση πόρων είναι πιο αποδοτικό (κατά μέσο όρο). Το συγκεκριμένο σύνολο δεδομένων εμφανίζει αρκετές δυσκολίες στη μελέτη του. Πρόκειται για ένα αρκετά πολύπλοκο σύνολο δεδομένων (περισσότερες λεπτομέρειες για τα χαρακτηριστικά του δίνονται στο Παράρτημα Β στο τέλος της διπλωματικής), το οποίο απαιτεί όπως φαίνεται μεγάλο αριθμό κανόνων για να λειτουργήσει σωστά και αποδοτικά. Στις πειραματικές δοκιμές που κάναμε, προκειμένου να περιορίσουμε τη χρονική διάρκεια εκτέλεσης και να κρατήσουμε σε ανεκτά επίπεδα τον αριθμό των προστιθέμενων κόμβων κανόνων, βάλουμε στο σύστημα έναν περιορισμό για προσθήκη μέχρι 100 κανόνων. Όπως φαίνεται από την πέμπτη στήλη του Πίνακα 4.1, τόσο το vehicle, όσο και το ionosphere, πλησίασαν αρκετά αυτό το όριο των εκατό κανόνων. Επιπλέον, όπως διαπιστώνεται από την πρώτη στήλη του Πίνακα 4.3, ο αριθμός των «σημαντικών» κανόνων, με βάση το κριτήριο που τέθηκε ($\Phi_i > 1$), είναι αρκετά υψηλός για το σύνολο δεδομένων Vehicle. Κάτι άλλο που μπορούμε επίσης να παρατηρήσουμε, είναι ότι το σύνολο δεδομένων Vehicle, χρειάστηκε κατά πολύ μεγάλο ποσοστό (87,96%) να χρησιμοποιήσει και το δευτερεύον υποσύνολο κανόνων (δηλαδή όλους τους κανόνες που προστέθηκαν στο σύστημα κατά την εκπαίδευση),

γεγονός που εμείς θέλουμε να αποφύγουμε και δεν εξυπηρετεί καθόλου τις επιδιώξεις μας. Αυτό που πρέπει βέβαια να σημειωθεί, είναι ότι το συγκεκριμένο υποσύνολο κανόνων είναι αρκετά πολύπλοκο και απαιτητικό. Περιλαμβάνει έτσι κι αλλιώς μεγάλο αριθμό προτύπων (846) με αρκετά χαρακτηριστικά (18), ενώ και ως είδος πληροφορίας τα δεδομένα φαίνονται σύνθετα (σιλουέτες οχημάτων). Επομένως όλα αυτά τα συμπεράσματα για τον αριθμό των κανόνων και τον επιπλέον χρόνο εκτέλεσης ήταν αναμενόμενα.

Ως επόμενο βήμα για περαιτέρω έρευνα μεταβάλαμε τα κριτήρια για την προσθήκη κανόνων στο σύστημα, δηλαδή το ϵ και το δ , ώστε να διαπιστώσουμε πώς αλλιώς μπορεί να γίνει διαχείριση της δομής του δικτύου μας. Πέρα από τη μεταβολή των τιμών των ϵ και δ , ώστε να ρυθμιστεί η «αυστηρότητα» των κριτηρίων για την εισαγωγή νέων κανόνων, εφαρμόσαμε και άλλο ένα κριτήριο, το ϵ_2 (ίδιας μορφής με το ϵ) το οποίο ορίζεται ως εξής:

$$\epsilon_2 = \frac{\text{διαφορά όλων των εξόδων}}{\text{αριθμό εξόδων}} = \frac{d_{\text{out}} - y_{\text{out}}}{\text{number_of_outputs}} \quad 4.1$$

όπου d_{out} είναι η επιθυμητή έξοδος, y_{out} η πραγματική έξοδος του συστήματος και number_of_outputs είναι ο αριθμός των εξόδων.

Συνδυάζοντας το κριτήριο αυτό με το δ , παίρνουμε τα αποτελέσματα που φαίνονται στον Πίνακα 4.5. Οι δοκιμές έγιναν με το σύνολο δεδομένων Pima Indians, και τα αποτελέσματα αντιπροσωπεύουν τους μέσους όρους 10 δοκιμών.

Σύνολο Δεδομένων	Αρχικός αριθμός κανόνων	ϵ_2 (M.O)	δ (M.O)	Τελικός αριθμός κανόνων (M.O)	Απόδοση Εκπαίδευσης (%,M.O)
Pima Indians	2	0,415	0,41	2,3	75,16

Πίνακας 4.5: Αποτελέσματα πειραμάτων για την εκπαίδευση του συστήματος με κριτήρια προσθήκης κανόνων ϵ_2 και δ

Σύνολο Δεδομένων	Απόδοση Προσομοίωσης (%, M.O)	Χρόνος εκτέλεσης (sec, M.O)
Pima Indians	76,77	0,5145

Πίνακας 4.6: Αποτελέσματα απλής προσομοίωσης εκπαιδευμένου συστήματος με κριτήρια ϵ_2 και δ

Σύνολο Δεδομένων	Πλήθος κανόνων με $\Phi_i > 1$	Χρόνος εκτέλεσης (sec, M.O)	Κατώφλι εμπιστοσύνης	Απόδοση Συστήματος (% , M.O)	Κύριο υποσύνολο κανόνων (% , M.O)	Και δευτερεύον υποσύνολο (% , M.O)
Pima Indians	1,4	0,44	0,5	72,134	53,73	46,27

Πίνακας 4.7: Αποτελέσματα προσομοίωσης εκπαιδευμένου συστήματος με ιεράρχηση κανόνων με κριτήριο την εναισθησία και κριτήρια προσθήκης κανόνων δ και ε_2 .

Σύνολο Δεδομένων	Απόδοση Προσομοίωσης (% , M.O)	Απόδοση Συστήματος (% , M.O)	Χρόνος εκτέλεσης1 (sec, M.O)	Χρόνος εκτέλεσης2 (sec, M.O)
Pima Indians	76,77	72,134	0,5145	0,44

Πίνακας 4.8: Συγκριτικός πίνακας για τις δυο προσομοιώσεις.

Τα αποτελέσματα που προκύπτουν για το νέο κριτήριο προσθήκης κανόνων είναι παρόμοια με τα προηγούμενα. Έχουμε συνολικά πιο αποδοτικό δίκτυο, με παρόμοια αριθμητικά απόδοση, αλλά με καλύτερο χρόνο εκτέλεσης και με μεγάλο ποσοστό χρησιμοποίησης μόνο του κύριου υποσυνόλου κανόνων.

Αυτό που μπορούμε να παρατηρήσουμε και από τα 10 πειράματα που εκτελέσαμε με κριτήρια τα δ και ε_2 , είναι ότι πρόκειται για «αυστηρά» κριτήρια, ακόμα και αν οι τιμές τους μεταβάλλονται προκειμένου να γίνουν πιο ελαστικά. Αυτό πρακτικά σημαίνει ότι δεν προστίθενται εύκολα κανόνες στο σύστημά μας, με αποτέλεσμα ο τελικά αριθμός κανόνων να είναι αρκετά μικρός, ελάχιστα αυξημένος από τον αρχικό αριθμό κανόνων που θέσαμε εμείς στο δίκτυο. Συγκεκριμένα, για τα δικά μας πειράματα, με το συγκεκριμένο σύνολο δεδομένων, το δίκτυο, ακόμα και με πολύ «ελαστικές» τιμές για τα δ και ε_2 , προσθέτει το πολύ δύο κρυφούς κόμβους- κανόνες, με αποτέλεσμα ο τελικός αριθμός κανόνων να μην ξεπερνάει σε καμία μέτρηση τους 4. Ο αριθμός αυτός είναι μικρότερος από αυτόν που έχουμε από τα προηγούμενα πειράματα για το ίδιο σύνολο δεδομένων (5,5 κατά μέσο όρο, όπως φαίνεται στον Πίνακα 4.1), ενώ αντίστοιχα το ποσοστό απόδοσης του συστήματος είναι σχεδόν το ίδιο, και λίγο αυξημένο είναι το ποσοστό χρησιμοποίησης όλων των κανόνων, κύριου και δευτερεύοντος υποσυνόλου. Αυτό σημαίνει ότι το δίκτυο προσθέτει λιγότερους κανόνες, άρα χρησιμοποιεί και λιγότερους κανόνες, ακόμα και όταν τους

Κεφάλαιο 4: Πειραματική Μελέτη

χρησιμοποιεί όλους, με το ίδιο σχεδόν ποσοστό απόδοσης. Προκύπτει δηλαδή ένα οικονομικότερο σύστημα, με μικρότερη κατανάλωση πόρων και την ίδια απόδοση.

Στη συνέχεια πραγματοποιήσαμε δοκιμές με μόνο κριτήριο το ϵ_2 και όχι το δ . Μετά από πέντε δοκιμές τα αποτελέσματα που προέκυψαν φαίνονται στους Πίνακες 4.9 έως 4.12:

Σύνολο Δεδομένων	Αρχικός αριθμός κανόνων	ϵ_2 (M.O)	Τελικός αριθμός κανόνων (M.O)	Απόδοση Εκπαίδευσης (% , M.O)
Pima Indians	2	0,3875	2,5	75,96

Πίνακας 4.9: Αποτελέσματα πειραμάτων για την εκπαίδευση του συστήματος με κριτήριο προσθήκης κανόνων ϵ_2

Σύνολο Δεδομένων	Απόδοση Προσομοίωσης (% , M.O)	Χρόνος εκτέλεσης (sec, AVG)
Pima Indians	77,08	0,56

Πίνακας 4.10: Αποτελέσματα απλής προσομοίωσης εκπαιδευμένου συστήματος με κριτήριο ϵ_2

Σύνολο Δεδομένων	Πλήθος κανόνων με $\Phi_i > 1$	Χρόνος εκτέλεσης (sec, M.O)	Κατώφλι εμπιστοσύνης	Απόδοση Συστήματος (% , M.O)	Κύριο υποσύνολο κανόνων (% , M.O)	Και δευτερεύον υποσύνολο (% , M.O)
Pima Indians	1,75	0,511	0,5	74,25	25	75

Πίνακας 4.11: Αποτελέσματα προσομοίωσης εκπαιδευμένου συστήματος με ιεράρχηση κανόνων με κριτήριο την εναισθησία και κριτήριο προσθήκης κανόνων ϵ_2 .

Σύνολο Δεδομένων	Απόδοση Προσομοίωσης (%, M.O)	Απόδοση Συστήματος (%, M.O)	Χρόνος εκτέλεσης1 (sec, M.O)	Χρόνος εκτέλεσης2 (sec, M.O)
Pima Indians	77,08	74,25	0,56	0,511

Πίνακας 4.12: Συγκριτικός πίνακας για τις δυο προσομοιώσεις

Όπως φαίνεται από τους πίνακες 4.9 έως 4.12, τα αποτελέσματα από τις δοκιμές με μοναδικό κριτήριο το ε_2 δε διαφέρουν αισθητά από τα αντίστοιχα για τις δοκιμές και με τα δύο κριτήρια. Αυτό μας οδηγεί στο συμπέρασμα ότι το ε_2 ως κριτήριο είναι αρκετά ισχυρό και αυστηρό, και έχει καθοριστική σημασία για την προσθήκη ή όχι κανόνων στο δίκτυό μας.

Πέρα από τη σύγκριση των μεθόδων που χρησιμοποιήσαμε για τη διαχείριση της δομής του δικτύου, από τις πειραματικές δοκιμές που εκτελέσαμε, μπορούμε να βγάλουμε χρήσιμα συμπεράσματα για τις παραμέτρους με τις οποίες αρχικοποιούμε το δίκτυο, ώστε να είναι όσο γίνεται πιο αποτελεσματικό.

Από τις δοκιμές που κάναμε με τα διάφορα σύνολα δεδομένων, παρατηρούμε ότι σε κάθε σύνολο, ο τελικός αριθμός κανόνων δε διαφοροποιείται σημαντικά, αν διατηρήσουμε τις υπόλοιπες παραμέτρους του δικτύου μας σταθερές. Αυτό μας δίνει τη δυνατότητα να θέσουμε ένα όριο στο δίκτυο για το πλήθος των κανόνων που μπορεί να προσθέσει, γνωρίζοντας ότι είναι σχεδόν απίθανο να χρειαστεί να προσθέσει περισσότερους. Όπως αναφέρθηκε και παραπάνω, στις δοκιμές μας θέσαμε τον περιορισμό της προσθήκης μέχρι 100 κανόνων στο δίκτυο, προκειμένου να αποφύγουμε κάποια φαινόμενα προσθήκης ακόμα και 800 (!) κανόνων, τα οποία παρατηρήθηκαν σε κάποιες περιπτώσεις με πολύπλοκα σύνολα δεδομένων (όπως το vehicle). Μετά το τέλος των δοκιμών, και έχοντας παρατηρήσει τη συμπεριφορά του δικτύου για 10 τουλάχιστον επαναλήψεις, μπορούμε να βγάλουμε συμπεράσματα για τον αριθμό των κανόνων που χρειάζεται κατά μέσο όρο το δίκτυο για να λειτουργήσει ικανοποιητικά. Για το συμπέρασμα αυτό, μπορούμε να παρατηρήσουμε τον αριθμό των κανόνων που καταλήγει να έχει το δίκτυο μετά το τέλος της εκπαίδευσης, αλλά και το πλήθος των κανόνων που θεωρεί το δίκτυο σημαντικούς και τους τοποθετεί στο κύριο υποσύνολο (έχουν δηλαδή ισχύ ενεργοποίησης μεγαλύτερη της μονάδας). Επιπλέον, σημαντική πληροφορία είναι και το ποσοστό με το οποίο το σύστημα χρησιμοποιεί το κύριο ή και το δευτερεύον υποσύνολο κανόνων, δηλαδή όλους τους κανόνες που καταλήγει να έχει το δίκτυο μετά την εκπαίδευση.

Για παράδειγμα, με βάση την πέμπτη στήλη του Πίνακα 4.1, όπου φαίνεται ο τελικός αριθμός κανόνων, μπορούμε να πούμε ότι για το σύνολο δεδομένων Breast Cancer, όπως και για το Pima Indians, ο τελικός αριθμός των κανόνων δεν ξεπερνάει τους 20 (14.5 και 5.5

κανόνες αντίστοιχα). Επομένως μπορούμε ως όριο για την προσθήκη κανόνων να θέσουμε το 50, αφήνοντας μάλιστα ένα μεγάλο περιθώριο ασφαλείας. Επίσης μπορούμε να βγάλουμε και ένα σχετικό συμπέρασμα και για τον αριθμό των κανόνων με τον οποίο θα αρχικοποιήσουμε το δίκτυο. Από την άλλη τα σύνολα δεδομένων Ionosphere και Vehicle, με 80,3 και 68,46 τελικό αριθμό κανόνων αντίστοιχα κατά μέσο όρο, δε μας επιτρέπουν να κατεβάσουμε το όριο των κρυφών κόμβων που μπορεί να προσθέσει το σύστημα κάτω από το 100, δεδομένης και της πολυπλοκότητας που εμφανίζουν τα δυο αυτά σύνολα δεδομένων και ιδιαίτερα το vehicle. Πρόκειται για δεδομένα με πολλά χαρακτηριστικά που απαιτούν μεγάλο αριθμό κανόνων και ίσως και μεγάλο αριθμό εποχών, προκειμένου να εκπαιδευτεί το δίκτυο σωστά και να λειτουργήσει ικανοποιητικά.

Παρακάτω δίνονται κάποια διαγράμματα που απεικονίζουν σχέσεις ανάμεσα σε βασικά μεγέθη της διαδικασίας, για κάθε σύνολο δεδομένων.

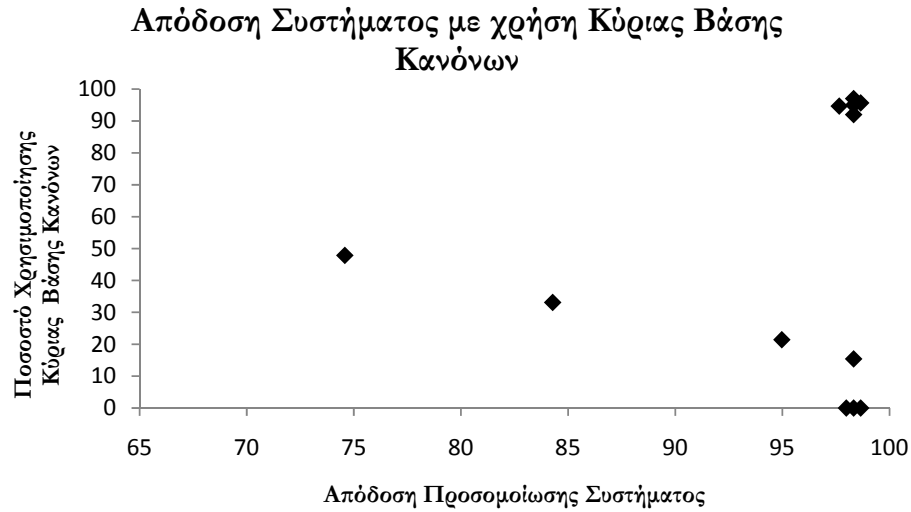
Για το σύνολο δεδομένων **Breast Cancer**:



Σχήμα 4.1: Απόδοση εκπαιδευμένου συστήματος για το σύνολο Breast Cancer

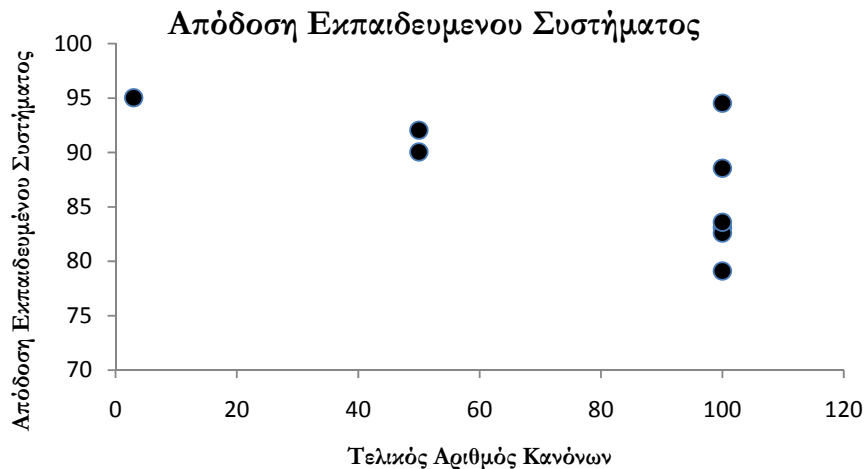


Σχήμα 4.2: Πλήθος κανόνων με ισχύ ενεργοποίησης > 1 για το σύνολο Breast Cancer

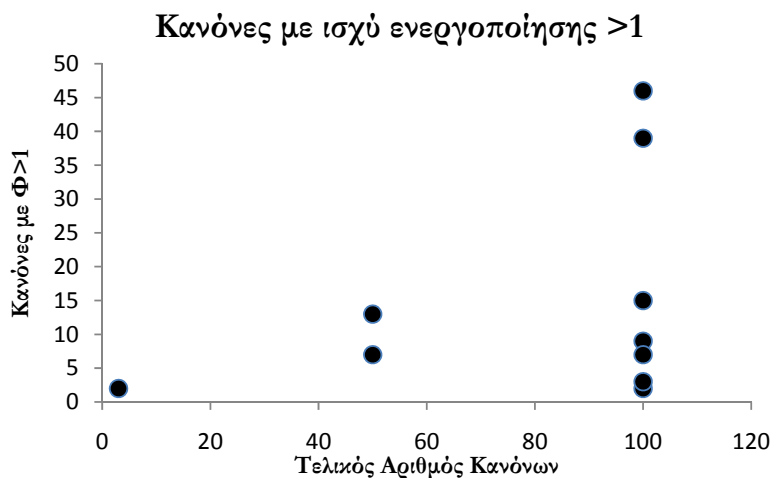


Σχήμα 4.3: Απόδοση συστήματος μόνο με την κύρια βάση κανόνων για το σύνολο Breast Cancer

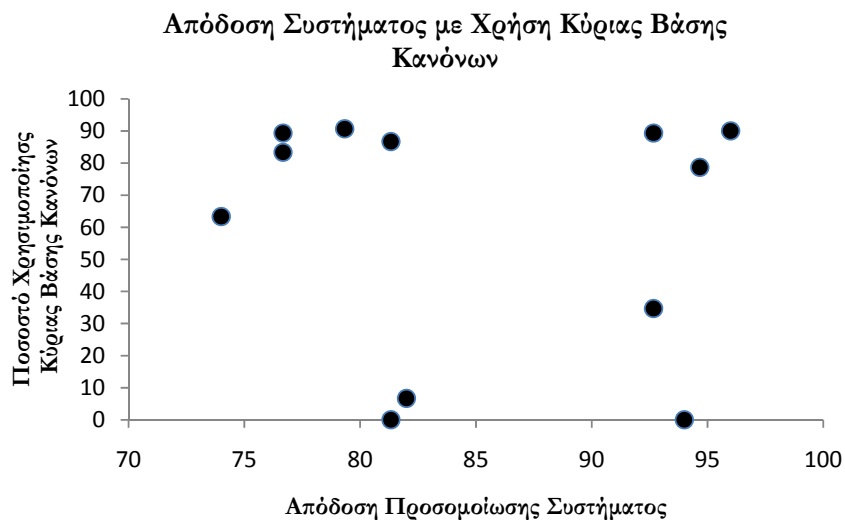
Για το σύνολο δεδομένων **Ionosphere**:



Σχήμα 4.4: Απόδοση εκπαιδευμένου συστήματος για το σύνολο Ionosphere

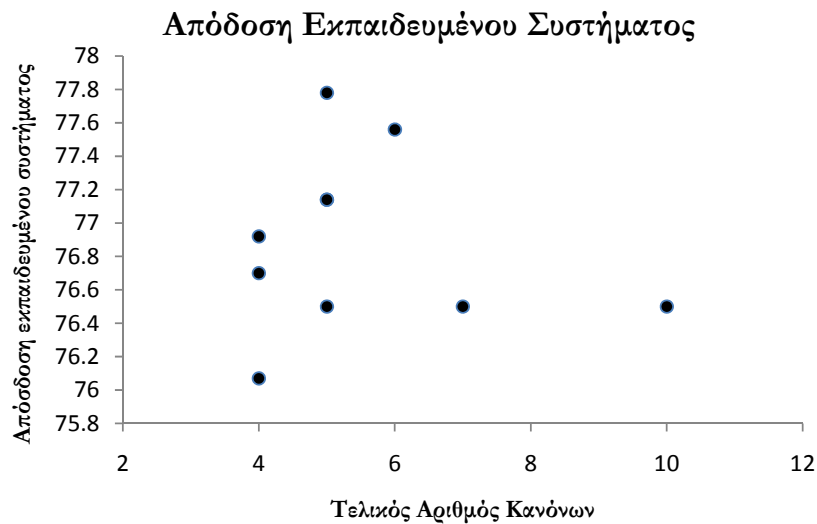


Σχήμα 4.5: Πλήθος κανόνων με ισχύ ενεργοποίησης >1 για το σύνολο Ionosphere

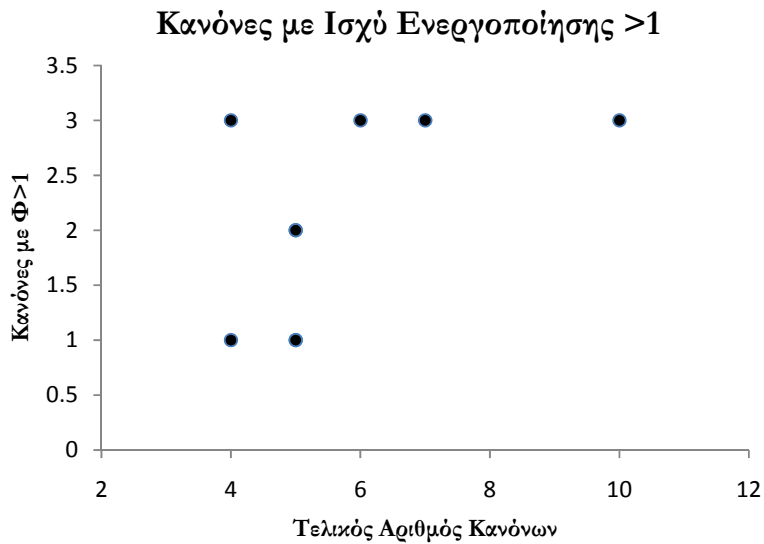


Σχήμα 4.6: Απόδοση συστήματος μόνο με την κύρια βάση κανόνων για το σύνολο Ionosphere

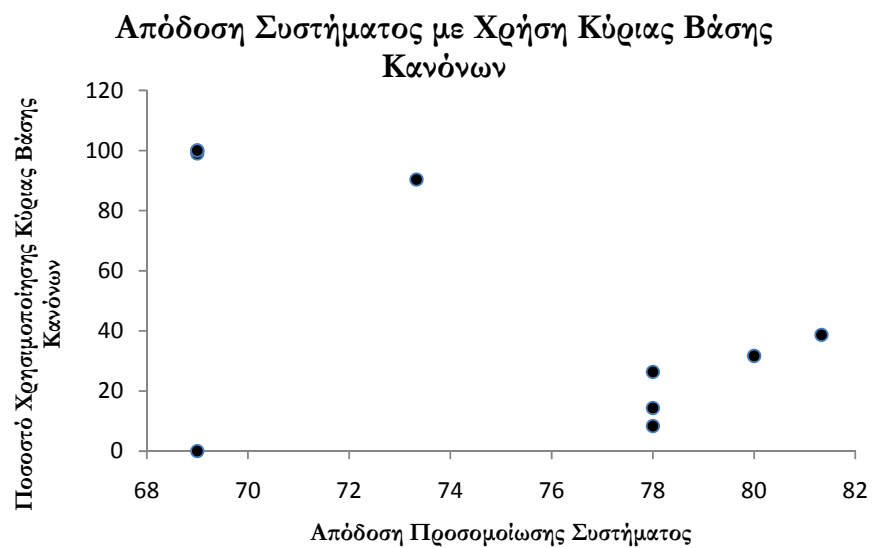
Για το σύνολο δεδομένων **Pima Indians**:



Σχήμα 4.7: Απόδοση εκπαιδευμένου συστήματος για το σύνολο Pima Indians

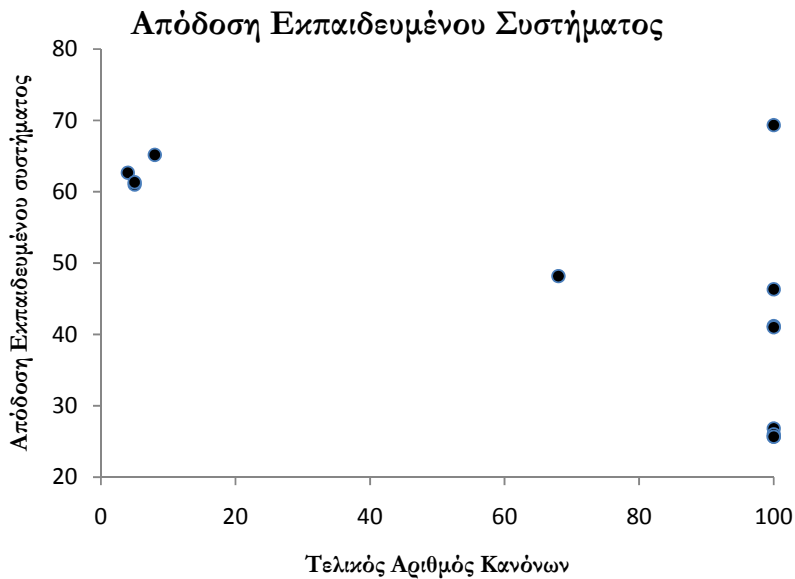


Σχήμα 4.8: Πλήθος κανόνων με ισχύ ενεργοποίησης >1 για το σύνολο Pima Indians

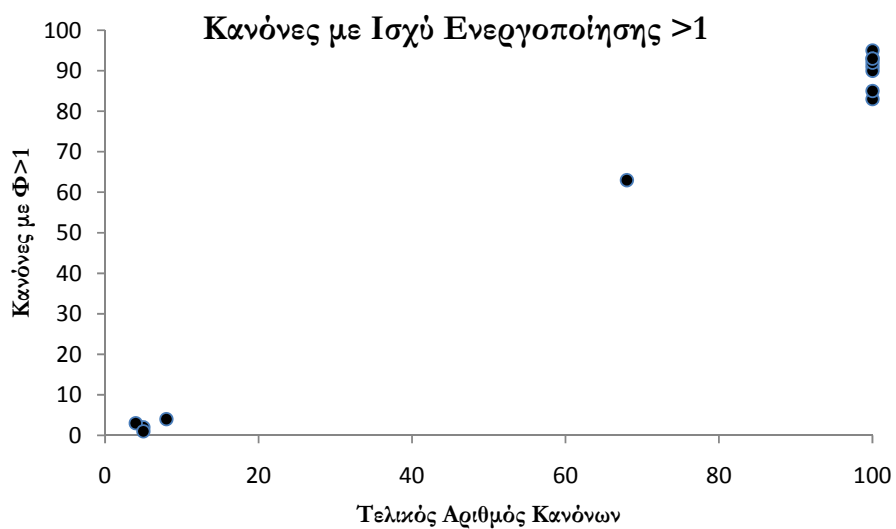


Σχήμα 4.9: Απόδοση συστήματος μόνο με την κύρια βάση κανόνων για το σύνολο Pima Indians

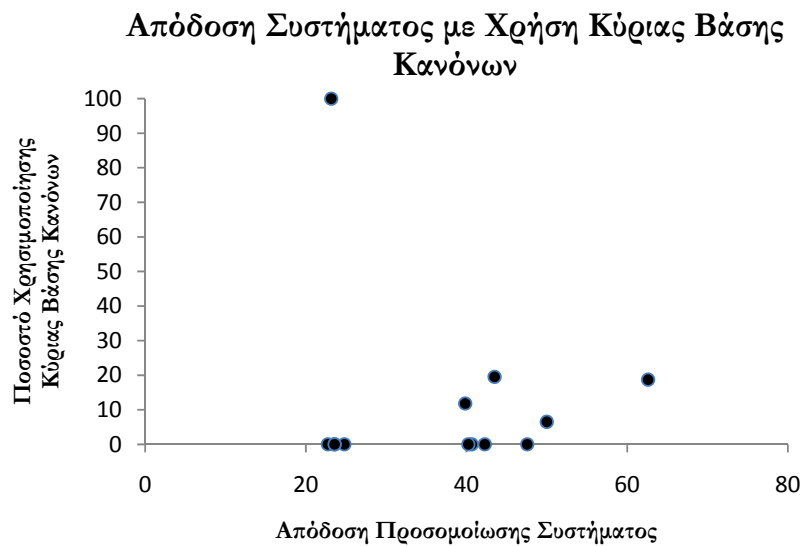
Για το σύνολο δεδομένων **Vehicle**:



Σχήμα 4.10: Απόδοση εκπαιδευμένου συστήματος για το σύνολο Vehicle



Σχήμα 4.11: Πλήθος κανόνων με ισχύ ενεργοποίησης >1 για το σύνολο Vehicle



Σχήμα 4.12: Απόδοση συστήματος μόνο με την κύρια βάση κανόνων για το σύνολο Vehicle

Κεφάλαιο 5: Συμπεράσματα– Μελλοντικές επεκτάσεις

5.1 Συμπεράσματα

Έχοντας πλέον ολοκληρώσει τις πειραματικές μετρήσεις και δοκιμές και έχοντας στη διάθεσή μας αρκετά πειραματικά αποτελέσματα για διαφορετικά σύνολα δεδομένων, μπορούμε πλέον να αποφανθούμε για τη χρησιμότητα και αποτελεσματικότητα των μεθόδων που εφαρμόσαμε και δοκιμάσαμε. Ήδη από το αρχικό στάδιο προσπαθήσαμε να δοκιμάσουμε να διαχειριστούμε δυναμικά τη δομή ενός δικτύου και να ρυθμίσουμε την κατανάλωση των πόρων που χρειάζεται και χρησιμοποιεί, προκειμένου να καταλήξουμε σε ένα δίκτυο που θα έχει την ίδια ή και καλύτερη απόδοση, ενώ συγχρόνως θα είναι και πιο οικονομικό, με χαμηλότερο λειτουργικό και σχεδιαστικό κόστος.

Αυτό που έπρεπε να διαπιστώσουμε αρχικά ήταν το αν η ιδέα για τη δυναμική επιλογή και διαχείριση πόρων και δομής, είναι εφικτή και αποδοτικά εφαρμόσιμη στον τομέα των νευροασφών συστημάτων. Για να το διαπιστώσουμε αυτό χρησιμοποιήσαμε μια μεθοδολογία αξιολόγησης και επιλογής κανόνων, η οποία στηρίζεται σε ένα πολύ γνωστό και εύχρηστο εργαλείο της στατιστικής, την ανάλυση ευαισθησίας. Φτιάξαμε έτσι μια νέα μέθοδο για την ιεράρχηση των κανόνων ενός δικτύου, η οποία ως διαδικασία ήταν αρκετά

απλή και κατανοητή, αλλά και εύκολα εφαρμόσιμη, αφού η ανάλυση ευαισθησίας είναι μια διαδομένη και εύκολη πρακτική μέθοδος. Τη μέθοδο αυτή δοκιμάσαμε να την ενσωματώσουμε σε ένα μοντέλο νευροασαφούς συστήματος που ήδη παρείχε τη δυνατότητα για δυναμική διαχείριση πόρων, το ARANFIS. Δεν ακουόν όμως οι θεωρητικές υποθέσεις. Για να επιβεβαιωθεί η χρησιμότητα των προτεινόμενων μεθόδων και να μπορέσουν αυτές οι νέες προτάσεις να εφαρμοστούν και πρακτικά, έπρεπε να υποστηρίξονται από τα αντίστοιχα πειραματικά αποτελέσματα. Το επόμενο, λοιπόν, βήμα ήταν μια σειρά από πειράματα και δοκιμές σε έτοιμα σύνολα δεδομένων, που προσφέρονται για πειραματικές δοκιμές. Προκειμένου να είναι, έγκυρα τα αποτελέσματα, δεν αρκεστήκαμε σε μικρό αριθμό δοκιμών ή σε λίγες τιμές των εμπλεκόμενων παραμέτρων, αλλά δοκιμάσαμε να μεταβάλλουμε τις τιμές στα διάφορα μεγέθη και να επαναλάβουμε τα πειράματα αρκετές φορές (το λιγότερο 10), ώστε να έχουμε μέσους όρους αντιπροσωπευτικούς της διαδικασίας και να μπορούμε με ασφάλεια να συγκρίνουμε τα αποτελέσματα.

Όπως φάνηκε και από τα αποτελέσματα των δοκιμών που παραθέσαμε στο κεφάλαιο 4, το νέο σύστημα, δηλαδή το μοντέλο του ARANFIS σε συνδυασμό με την αξιολόγηση και ιεράρχηση κανόνων με βάση την ευαισθησία, ικανοποιεί σε μεγάλο βαθμό τους στόχους που θέσαμε στην αρχή της διπλωματικής αυτής. Προσφέρει δηλαδή ένα σύνθετο μοντέλο με προσαρμοζόμενη ανάθεση πόρων και ταυτόχρονα δυναμική διαχείριση της δομής του συστήματος κατά την εκπαίδευση, με την αξιολόγηση και δυναμική επιλογή των κανόνων που θα χρησιμοποιήσει τελικά. Τα αποτελέσματα, τόσο ως προς την απόδοση του συστήματος, όσο και ως προς το χρόνο εκτέλεσης και το ποσοστό των διαθέσιμων πόρων που χρησιμοποιούνται, ήταν ιδιαίτερα ευνοϊκά, και επιβεβαίωσαν σε μεγάλο βαθμό τις θεωρητικές υποθέσεις. Με το συνδυασμό που περιγράφουμε, καταφέραμε να δημιουργήσουμε ένα σύστημα με την ίδια απόδοση με τα υπάρχοντα συστήματα, αλλά με πολύ καλύτερα επίπεδα σχεδιαστικού και λειτουργικού κόστους: χρησιμοποίηση μικρότερου αριθμού κανόνων, σε μικρότερο χρονικό διάστημα εκτέλεσης. Επιπλέον, προέκυψαν συμπεράσματα σχετικά με τις παραμέτρους του συστήματος, που θα μπορούσαν να χρησιμοποιηθούν για την αρχικοποίησή του, ώστε να έχουμε ακόμα καλύτερα αποτελέσματα. Το σύστημά μας υπέδειξε ότι η προσαρμογή της δομής ενός δικτύου στα δεδομένα παρουσιάζει σημαντικά πλεονεκτήματα έναντι της κλασσικής τυχαίας ή στατικής αρχιτεκτονικής. Ταυτόχρονα μας δόθηκε η δυνατότητα να παρατηρήσουμε και κάποιες αδυναμίες των νευροασαφών συστημάτων που πρέπει να εξαλειφθούν, κυρίως όσον αφορά στη δυναμική διαχείριση και προσαρμογή της δομής και των πόρων τους στις ελάχιστες συνθήκες της κάθε εφαρμογής, ώστε να επιτευχθεί και η διαρκώς επιθυμητή προσαρμοστικότητα.

Το συνολικό, βέβαια συμπέρασμα που εξάγεται από το σύνολο της έρευνας για αυτήν τη διπλωματική, είναι πως τα προσαρμοστικά νευροασαφή δίκτυα γινομένου αποτελούν ένα σημαντικό και εύχρηστο εργαλείο παραγωγής αποφάσεων, το οποίο μπορεί να εφαρμοστεί για την επίλυση οποιουδήποτε προβλήματος, με τρόπο που ο άνθρωπος μπορεί να

κατανοήσει και να ελέγξει ως ένα βαθμό. Τα προβλήματα που εμφανίζονται κατά τη λειτουργία τους και αφορούν στην αργή επεξεργασία των δεδομένων και την εκθετική αύξηση της πολυπλοκότητας συναρτήσει των προτύπων εισόδου, αδυναμία που μας ανάγει στη λεγόμενη «κατάρρα της διάστασης», επιβάλλεται να ξεπεραστούν με τη χρήση νέων και σύγχρονων τεχνικών και τεχνολογιών. Σε αυτόν τον τομέα προσπαθήσαμε να προσφέρουμε κάποια νέα δεδομένα. Προκειμένου να αντιμετωπίσουμε, έστω και σε αρχικό στάδιο, την κατάρρα της διάστασης, προτεινάμε ένα δίκτυο που θα ελέγχει κατά τη λειτουργία του το μέγεθός του. Είναι ικανό να επεξεργαστεί και να διαχειριστεί τόσο τα δεδομένα εισόδου, όσο και την εσωτερική του δομή, έχοντας ταυτόχρονα τη μικρότερη απώλεια πληροφορίας. Μεταβάλλοντας δυναμικά τους πόρους που καταναλώνει, αποφεύγει να πέσει στις παγίδες που κρύβει μια στατική δομή και προσφέρει μεγάλη δυνατότητα προσαρμοστικότητας, καλύπτοντας μια από τις βασικότερες ανάγκες των σύγχρονων ευφυών συστημάτων. Επιπλέον, μειώνοντας τον αριθμό των πόρων που διαχειρίζεται και των συνάψεων που χρειάζεται για να λειτουργήσει (προσαρμοζόμενος αριθμός κόμβων κανόνων), συμβάλλει και στην επίτευξη ενός ακόμη δύσκολου αλλά κυρίαρχου στόχου της σύγχρονης επιστημονικής κοινότητας, αυτόν της εξοικονόμησης ενέργειας και πόρων, της μείωσης ενεργειακού, λειτουργικού και σχεδιαστικού κόστους, της ανάπτυξης μιας επιστήμης με οικολογική συνείδηση και περιβαλλοντική ευαισθησία.

5.2 Μελλοντικές Επεκτάσεις

Όπως αναφέρθηκε και στην εισαγωγή της παρούσας διπλωματικής, οι τομείς της περιβαλλοντικής ευφυΐας και των νευροασαφών συστημάτων, τομείς που σχετίζονται άμεσα μεταξύ τους, είναι από τους πλέον σύγχρονους και αναπτυσσόμενους στο χώρο της τεχνολογίας. Στην έρευνα αυτή προσπαθήσαμε να εφαρμόσουμε νέες ιδέες και μεθόδους, προκειμένου να εξαιλέψουμε κάποιες από τις υπάρχουσες αδυναμίες, κυρίως στον τομέα των νευροασαφών συστημάτων. Είναι βέβαια προφανές, ότι ο χώρος αυτός προσφέρει απεριόριστες ευκαιρίες για περαιτέρω έρευνα, και υπάρχει ακόμα ένα πλήθος θεμάτων που μπορούν να μελετηθούν, να αναπτυχθούν και να βελτιωθούν αισθητά.

Η διαχείριση δομής που προτεινάμε εμείς βασίζεται στην ανάλυση ευαισθησίας με κύριο πρωταγωνιστή την ισχύ ενεργοποίησης των κρυφών κόμβων κανόνων. Πριν φτάσουμε βέβαια στο επίπεδο των κρυφών κόμβων, μπορούμε να ασχοληθούμε πιο αναλυτικά με το επίπεδο των προτύπων εισόδου. Ακόμα και αν η προεπεξεργασία των δεδομένων εισόδου είναι ήδη κάτι συνηθισμένο που εφαρμόζεται ευρέως, αυτό δε σημαίνει ότι δεν επιδέχεται βελτίωση. Μια καλύτερη προεπεξεργασία και προετοιμασία των δεδομένων εισόδου, η οποία θα περιλαμβάνει την αξιολόγηση και διαλογή τους, μπορεί να απομονώσει τα πιο σημαντικά δεδομένα για το σύστημα, με την ελάχιστη όμως απώλεια πληροφορίας.

Ένα άλλο σημείο προς μελέτη, το οποίο προσπαθήσαμε να προσεγγίσουμε και στην παρούσα διπλωματική εργασία, είναι η καλύτερη αρχικοποίηση του δικτύου μας. Ένα βήμα προς αυτήν την κατεύθυνση είναι η χρήση του τελικού αριθμού κανόνων που προκύπτει από την εκπαίδευση ως ανώτατο όριο για τους κόμβους που προστίθενται. Πρόκειται όμως για μια αρχική πρόταση, η οποία έχει πολλά περιθώρια να μελετηθεί και να αναλυθεί πολύ περισσότερο. Ο τελικός αριθμός κανόνων δεν είναι το μόνο μέγεθος που μπορεί να συμβάλει σε μια καλύτερη αρχικοποίηση του δικτύου. Ακόμα και τα κριτήρια προσθήκης κανόνων ή το κατώφλι εμπιστοσύνης, είναι μεγέθη που μπορούν να ερευνηθούν περαιτέρω και να αποδειχθούν πιθανόν ιδιαίτερα χρήσιμα.

Τα κριτήρια προσθήκης νέων κόμβων κανόνων που προτείνουμε δεν είναι βέβαια τα μοναδικά. Οι συσχετισμοί που μπορούν να γίνουν μεταξύ των εμπλεκόμενων μεγεθών, αλλά και ο ρόλος των εισόδων, μπορούν να μελετηθούν και να δώσουν νέες ιδέες για κριτήρια προσθήκης κανόνων, τα οποία μπορούν να λειτουργήσουν είτε μόνα τους είτε σε συνδυασμό με αυτά που προτείνονται εδώ. Ακόμα και οι τιμές των κριτηρίων αυτών, παρόλο που μελετήθηκαν σε βάθος σε αυτήν τη διπλωματική, μπορούν να αποτελέσουν αντικείμενο για περαιτέρω έρευνα.

Προχωρώντας στο εσωτερικό του δικτύου, αφού έχουν μελετηθεί η προεπεξεργασία των δεδομένων και όσα αφορούν στην προθήκη νέων κόμβων κανόνων, επόμενο αντικείμενο για έρευνα είναι η καλύτερη διαχείριση και χρησιμοποίηση τόσο των δεδομένων εισόδου μετά την προεπεξεργασία, όσο και των κόμβων κανόνων που προστέθηκαν στο δίκτυο. Η λεγόμενη «κατάρρα της διάστασης» (“curse of dimensionality”), αποτελεί ένα από τα σημαντικότερα προβλήματα στον τομέα των νευροσασφών δικτύων. Νέες μέθοδοι και τεχνικές διαχείρισης της δομής και των πόρων ενός δικτύου είναι επιτακτική ανάγκη να εφαρμοστούν, ώστε να γίνουν βήματα σε αυτόν τον τομέα. Ο συνδυασμός ARANFIS και ανάλυσης ευαισθησίας -με τον τρόπο που εφαρμόζεται εδώ, βασισμένη στην ισχύ ενεργοποίησης των κανόνων- στρέφεται προς αυτήν την κατεύθυνση. Όμως τα περιθώρια για έρευνα και βελτίωση είναι τεράστια.

Παράρτημα Α:

Υπολογισμός μερικών παραγώγων για το μοντέλο ARANFIS

Ο υπολογισμός των μερικών παραγώγων που εμφανίζονται στις εξισώσεις 3.8 έως 3.12 του τρίτου κεφαλαίου δίνεται αναλυτικά παρακάτω. Για τον υπολογισμό αυτό θεωρήσαμε ότι το δίκτυο αποτελείται από n εισόδους, p εξόδους, και $q(t)$ κρυφούς κόμβους για την επανάληψη t (αφού το δίκτυο ARANFIS έχει προσαρμοζόμενη δομή ανάθεσης πόρων).

Για τη μερική παράγωγο του σφάλματος συναρτήσει των κέντρων των μετασυναπτικών βαρών θα έχουμε:

$$\frac{\partial e(t)}{\partial v_{jk}^c(t)} = -(d_k - y_k) \frac{z_j v_{jk}^\sigma}{\sum_{j=1}^{q(t)} z_j v_{jk}^\sigma} \quad \text{A.1}$$

όπου z_j η ισχύς ενεργοποίησης του κανόνα j που δίνεται από τη σχέση:

$$z_j = \prod_{i=1}^n E_{ij} \quad \text{A.2}$$

Και E_{ij} αμοιβαία σχέση εγκλεισμού η οποία υπολογίζεται από τη σχέση:

$$E_{ij} = E(s_i, w_{ij}) = \frac{C(s_i \cap w_{ij})}{C(s_i) + C(w_{ij}) - C(s_i \cap w_{ij})} \quad \text{A.3}$$

όπου s_i είναι το ασαφές σήμα εισόδου $S(x_i) = S(x_i) = (x_i^e, x_i^\sigma)$, και $C(A) = \int_{-\infty}^{+\infty} a(x) dx$ είναι

το μέγεθος του ασαφούς συνόλου A που περιγράφεται από τη συνάρτηση συμμετοχής $a(x)$ και επομένως $C(A \cap B)$ είναι η τομή των συναρτήσεων συμμετοχής των συνόλων A και B .

Για τη μερική παράγωγο του σφάλματος συναρτήσει της διασποράς των μετασυναπτικών βαρών θα έχουμε:

$$\frac{\partial e(t)}{\partial v_{jk}^\sigma(t)} = -(d_k - y_k) \frac{z_j v_{jk}^c \sum_{j=1}^{q(t)} z_j v_{jk}^\sigma - z_j \sum_{j=1}^{q(t)} z_j v_{jk}^c v_{jk}^\sigma}{\left(\sum_{j=1}^{q(t)} z_j v_{jk}^\sigma \right)^2} \quad \text{A.4}$$

όπου τα μεγέθη που εμφανίζονται έχουν εξηγηθεί και στις σχέσεις A.1 έως A.3.

Οι μερικές παράγωγοι του σφάλματος συναρτήσει των κέντρων και της διασποράς των προσυναπτικών βαρών, απαιτούν τον υπολογισμό και άλλων μερικών παραγώγων και είναι πιο σύνθετες στον υπολογισμό. Πιο συγκεκριμένα, η μερική παράγωγος του σφάλματος συναρτήσει των κέντρων των προσυναπτικών βαρών υπολογίζεται ως εξής:

$$\frac{\partial e}{\partial w_{ij}^c} = \sum_{k=1}^p -(d_k - y_k) \frac{\partial y_k}{\partial z_j} \frac{\partial z_j}{\partial E_{ij}} \frac{\partial E_{ij}}{\partial w_{ij}^c} \quad \text{A.5}$$

$$\frac{\partial e}{\partial w_{ij}^\sigma} = \sum_{k=1}^p -(d_k - y_k) \frac{\partial y_k}{\partial z_j} \frac{\partial z_j}{\partial E_{ij}} \frac{\partial E_{ij}}{\partial w_{ij}^\sigma} \quad \text{A.6}$$

Οι μερικές παράγωγοι του σφάλματος συναρτήσει των διασπορών των χαρακτηριστικών εισόδου, υπολογίζονται ως εξής:

$$\frac{\partial e(t)}{\partial x_i^\sigma(t)} = \sum_{j=1}^{q(t)} \sum_{k=1}^p -(d_k - y_k) \frac{\partial y_k}{\partial z_j} \frac{\partial z_j}{\partial E_{ij}} \frac{\partial E_{ij}}{\partial x_i^\sigma} \quad \text{A.7}$$

Όπου:

$$\frac{\partial y_k(t)}{\partial z_j(t)} = \frac{v_{jk}^\sigma (v_{jk}^c - y_k)}{\sum_{j=1}^{q(t)} z_j v_{jk}^\sigma} \quad \text{A.8}$$

Και

$$\frac{\partial z_j}{\partial E_{ij}} = \frac{\prod_{i=1}^n E_{ij}}{E_{ij}} \quad \text{A.9}$$

$$y_k(t) = \frac{\prod_{j=1}^{q(t)} z_j \mathcal{V}_{jk}^c \mathcal{V}_{jk}^s}{\prod_{j=1}^{q(t)} z_j \mathcal{V}_{jk}^s} \quad \text{A.10}$$

Παράρτημα Β:

Χαρακτηριστικά Προβλήματα Ταξινόμησης

Η βάση δεδομένων UCI αποτελεί ίσως τη μεγαλύτερη αποθήκη πειραματικών δεδομένων και οι περισσότερες ερευνητικές μελέτες στη βιβλιογραφία απευθύνονται εκεί για την επιλογή του κατάλληλου υλικού για τις δοκιμές και την αξιολόγηση των μεθόδων τους. Στη συγκεκριμένη διπλωματική εργασία χρησιμοποιήθηκαν τα σύνολα δεδομένων που αναφέρονται παρακάτω και τα οποία βρίσκονται διαθέσιμα στον ιστοχώρο:

www.ics.uci.edu/~mllearn/MLRepository.html

Σύνολο Δεδομένων Breast Cancer Wisconsin

Η συγκεκριμένη βάση δεδομένων για καρκίνο του μαστού, παρέχεται από το πανεπιστήμιο Wisconsin των ΗΠΑ και περιλαμβάνει την ταξινόμηση όγκων σε καλοήθεις και κακοήθεις, βάσει 9 χαρακτηριστικών που περιγράφουν τη φυσιολογία των κυττάρων. Το σύνολο δεδομένων περιλαμβάνει 699 πρότυπα.

Σύνολο Δεδομένων Ionosphere

Το σύνολο δεδομένων Ionosphere προέρχεται από ένα σύστημα ραντάρ στο Goose Bay του Labrador, και αποτελείται από 351 πρότυπα που περιγράφονται από 34 συνεχή χαρακτηριστικά και κατηγοριοποιούν τα ελεύθερα ηλεκτρόνια της ιονόσφαιρας σε καλά και κακά (2 κατηγορίες), ανάλογα με το αν σχηματίζουν κάποια δομή ή όχι. Από τα 351 πρότυπα, τα 200 χρησιμοποιούνται για εκπαίδευση του συστήματος (training), και τα υπόλοιπα 151 για αξιολόγηση (testing), όπως αναφέρεται συνήθως στη βιβλιογραφία.

Σύνολο Δεδομένων Pima Indians

Το σύνολο δεδομένων Pima Indians αποτελείται από 768 πρότυπα που αντιστοιχούν σε άτομα θηλυκού γένους με κληρονομικά χαρακτηριστικά της ινδιάνικης φυλής Pima. Τα 768 αυτά πρότυπα διακρίνονται από 8 βιολογικά αριθμητικά γνωρίσματα και χωρίζονται σε 2 κατηγορίες ανάλογα με το αν το άτομο εμφανίζει συμπτώματα διαβήτη, ή όχι.

Σύνολο Δεδομένων Vehicle

Το σύνολο Vehicle προέρχεται από το ινστιτούτο Turing της Γλασκόβης και περιλαμβάνει 846 σιλουέτες οχημάτων υπό διάφορες γωνίες, οι οποίες χωρίζονται σε 4 κατηγορίες: OPEL, SAAB, BUS, VAN. Η κάθε σιλουέτα περιγράφεται από 18 χαρακτηριστικά.

Βιβλιογραφία

- [1] Abowd, G., Det, A., Orr, R., and Brotherton, J. Context- awareness in wearable and ubiquitous computing. In *Proc. of the 1st Int. Symposium on Wearable Computers- ISWC* (1997), pp. 179- 180.
- [2] Balomenos, T., Pertselakis, M., McMahon, E., and Tsapatsoulis, N.. On the estimation of car driver profiles. In *Proc. of Int. Conf. "From Scientific Computing to Computational Engineering"- SCCE* (2004)
- [3] Cai, L., and Kwan, H. Fuzzy classifications using fuzzy inference networks. *IEEE Trans. on Systems, Man and Cybernetics* 28 (1998), 334- 337.
- [4] Chinneck, J. W., Practical Optimization: A Gentle Introduction, (<http://www.sce.carleton.ca/faculty/chinneck/po.html>)
- [5] Cordella, L. P., Foggia, P., Sansone, C., Tortorella, F., and Vento, M. Reliability parameters to improve combination strategies in multi-expert systems. *Pattern Analysis and Application* 2 (1999), 205- 214.
- [6] Engelbrecht, P., Cloete, I., and Zurada, J. M., Determining the significance of input parameters using sensitivity analysis. In *Int. Workshop on Artificial Neural Networks, "From Natural to Artificial Neural Computing"* (1995), J. Miral and F. Sandoval, Eds., vol. 930 of LNCS, pp.382- 388.
- [7] Howes, P., and Crook, N. Using input parameter influences to support the decisions of feedforward neural networks. *Neurocomputing* 24, 1-3 (1999), 191- 206.
- [8] Ioannou, S., Raouzaïou, A., Karpouzis, K., Pertselakis, M., Tsapatsoulis, N., and Kollias, S. Adaptive rule- based facial expression recognition. In *Proc. of 3rd Hellenic Conf. on Artificial Intelligence- SETN* (2004), vol. 3025 of *LNAI*, Springer- Verlag, pp. 466- 475.
- [9] Jang, J. S. R. ANFIS: Adaptive-network- based fuzzy inference system. *IEEE Trans. on Systems, Man & Cybernetics* 23, 1 (1993), 665- 685.
- [10] J. S. R. Jang, C. T. Sun, and E. Mizutani. *Neuro- Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*. Prentice Hall, 1997.

- [11] Jin, Y. Fuzzy modeling of high dimensional systems: Complexity reduction and interpretability improvement. *IEEE Trans. on Fuzzy Systems* 8 (2000), 212- 221.
- [12] Kim, J., and Kasabov, N. HyFIS: adaptive neuro- fuzzy inference systems and their application to nonlinear dynamical systems. *Neural Networks* 12, 9 (1999), vol. 3, pp. 1406- 1411.
- [13] Lee, R. S. T. *Fuzzy- Neuro Approach to Agent Applications: From the A.I Perspective to Modern Ontology*. Springer Series in Agent Technologies. Springer, 2006.
- [14] Levin, A. U., Leen, T. K., and Moody, J. E. Fast pruning using principal components. In *Advances in Neural Information Processing Systems* 6 (1994), J. D. Cowan, G. Tesauro, and J. Alspector, eds., pp. 35-42. Morgan Kaufmann, San Mateo, CA.
- [15] Lin, C. T., and Lee, C. S. G. *Neural fuzzy systems: a neuro- fuzzy synergism to intelligent systems*. Prentice Hall, 1996.
- [16] Mitra, S., and Hayashi, Y. Neuro- fuzzy rule generation: Survey in soft computing framework. *IEEE Trans. in Neural Networks* 11 (2000), 748- 768.
- [17] Moreira, M., and Fiesler, E. Neural Networks with adaptive learning rate and momentum terms. Tech. Rep. 95- 04, IDIAP, Martigny, Switzerland, 1995.
- [18] Nauck, D., and Kruse, R. A neuro- fuzzy method to learn fuzzy classification rules from data. *Fuzzy Sets and Systems* 89 (1997), 277- 288.
- [19] Pal, N. R., and Pal, T. On rule pruning using fuzzy neural networks. *Fuzzy Sets and systems* 106 (1999), 335- 347.
- [20] Paul, S., and Kumar, S. Subsethood- product fuzzy neural inference system (SuPFuNIS). *IEEE Trans. on Neural Networks* 13, 3 (2002), 578- 599.
- [21] Pertselakis, M., Tsapatsoulis, N., Kollias, S., and Stafylopatis, A. An adaptive resource allocating neural fuzzy inference system. In *Proc. IEEE Intelligent Systems Application to Power System- ISAP* (2003).
- [22] Platt, J. A resource- allocating network for function interpolation. *Neural Computation* 3, 2 (1991), 213- 225.

- [23] Solodov, M. V., and Svaiter, B. F., A comparison of rates of convergence of two inexact proximal point algorithms. In *Nonlinear optimisation and related topics* (2000), G. D. Pillo, and F. Giannesi, Eds., vol. 36 of *Applied Optimization*, Kluwer Academic Publishers, pp. 415- 427.
- [24] Velayutham, C. S., and Kumar, S. Assymmetric Subsethood- product fuzzy neural inference system (ASuPFuNIS). *IEEE Trans. on Neural Networks* 16, 1 (2005), 160- 174.
- [25] Velayutham, C. S, Paul, S., Kumar, S. Evolutionary subsethood product fuzzy neural network. In *Proc. of Int. Conf. on Fuzzy Systems- AFSS* (2002), N. R. Paul and M. Sugeno, Eds., vol. 2275 of *LNCS*, Springer- Verlag, pp.274- 280.
- [26] Yager, R. R., and Filev, D. P. *Essentials of Fuzzy Modelling and Control*. John Wiley & Sons, New York, 1994.
- [27] Zadeh, L.A. Fuzzy logic, neural networks, and soft computing. *Communications of ACM* 37 (1994), 77- 84.
- [28] Zadeh, L. A. Fuzzy sets. *Information & Control* 8 (1965), 338- 353.
- [29] Zeng, X., Yeung, D. S., and Sun, X. Sensitivity Analysis of Multilayer Perceptron to Input Perturbation. In *Proc. of IEEE Systems, Man, and Cybernetics, Int. Conf.* (2000), pp. 2509- 2514.
- [30] Zurada, J. M., Malinowski, A., and Cloete, I.. Sensitivity analysis for minimization of input data dimension for feedforward neural network. In *Proc. of IEEE Symposium on Circuits and Systems* (1994), pp. 447- 450.

Βιογραφικό Σημείωμα

Προσωπικές Πληροφορίες

Χρονολογία Γέννησης: 08/ 5 /1984

Τόπος γέννησης: ΑΘΗΝΑ

Σπουδές

2002 - 2008 Σχολή Ηλεκτρολόγων Μηχανικών και
Μηχανικών Υπολογιστών,
Ε.Μ.Π. ΑΘΗΝΑ
Κατεύθυνση Πληροφορικής

1996 – 2002 Ελληνογαλλική Σχολή Ουρσουλινών,
Ν. Ψυχικό, Αθήνα
Βαθμός Απολυτηρίου: 19,6

Γλώσσες

Αγγλικά : Certificate of Proficiency in English

Επάρκεια Διδασκαλίας της Αγγλικής Γλώσσας

Γαλλικά : Diplôme d' études françaises (2^o degré) (Sorbonne II)
Option : Littérature

Diplôme Approfondi de Langue Française (D.A.L.F.)
Option : Littérature

Επάρκεια Διδασκαλίας της Γαλλικής Γλώσσας

Συμμετοχές σε Εκπαιδευτικά Προγράμματα

2002 2^o Συμπόσιο Εταιρικών Σχολείων UNESCO με θέμα
«Θέματα Βιοηθικής στο σύγχρονο κόσμο»

2001 1^o Συμπόσιο Εταιρικών Σχολείων UNESCO με θέμα
«Η Ειρήνη στον 21^o αιώνα»

1999 Comenius-Δράση I
Ευρωπαϊκό πρόγραμμα με θέμα :
«Παγκόσμια πολιτιστική κληρονομιά που έχει
αναγνωριστεί από την Unesco στην Ευρώπη»
Συνεργασία με μαθητές από τη Γαλλία, το Βέλγιο, και την

Πορτογαλία στα πλαίσια συζητήσεων και καλλιτεχνικών δραστηριοτήτων – Συγγραφή Πρακτικών

Ικανότητες

Γλώσσες Προγραμματισμού: C, C++, Turbo Pascal, FORTRAN, Java, SQL, Matlab, Prolog, Assembly

Λειτουργικά Περιβάλλοντα: MS-Windows 98, 2000, Me, XP

Ανάπτυξη Πολυμεσικών Εφαρμογών: 3D Studio MAX R4, Corel Draw, Adobe Photoshop

Λοιπές Εφαρμογές: Internet Explorer, Outlook Express, Mozilla Firefox, Microsoft Word, Microsoft Excel, Microsoft PowerPoint, Microsoft Visio, PS-Spice

Ενδιαφέροντα/

**Λοιπές
Δραστηριότητες**

Συμμετοχή στους Ολυμπιακούς Αγώνες 2004, στην Αθήνα, ως Εξειδικευμένη Εθελόντρια στον Τομέα SP Services (Τομέας Εξυπηρέτησης Θεατών – Info Booth, ΟΑΚΑ)

Λογοτεχνία, Θέατρο, Κινηματογράφος, Μουσική, Διαδίκτυο