

Ατομική Διπλωματική Εργασία

ΑΣΑΦΕΙΣ ΓΝΩΣΤΙΚΟΙ ΧΑΡΤΕΣ

ΚΑΤΕΡΙΝΑ ΧΑΡΙΛΑΟΥ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΥΠΡΟΥ



ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

Μάιος 2011

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΥΠΡΟΥ

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

Ασαφείς Γνωστικοί Χάρτες

Κατερίνα Χαριλάου

Επιβλέπων Καθηγητής

Σχίζας Χρίστος

Η Ατομική Διπλωματική Εργασία υποβλήθηκε προς μερική εκπλήρωση των απαιτήσεων απόκτησης του πτυχίου Πληροφορικής του Τμήματος Πληροφορικής του Πανεπιστημίου Κύπρου

Μάιος 2011

Ευχαριστίες

Σε αυτό το σημείο θα ήθελα να ευχαριστήσω, τον επιβλέποντα καθηγητή μου Δρ. Χρίστο Σχίζα για την εμπιστοσύνη που μου επέδειξε με την ανάθεση του συγκεκριμένου θέματος αλλά και για την υποστήριξη, την καθοδήγηση και τις συμβουλές που μου προσέφερε καθ' όλη τη διάρκεια αυτής της Διπλωματικής Εργασίας.

Ιδιαίτερες ευχαριστίες, θα ήθελα να δώσω στην διδακτορική φοιτήτρια Μαρία Παπαϊωάννου, της οποίας τόσο η βοήθεια, η καθοδήγηση και η ψυχολογική υποστήριξη ήταν καθοριστική για την ολοκλήρωση της Διπλωματικής μου εργασίας.

Επίσης θα ήθελα να ευχαριστήσω τον Δρ. Νικόλα Χριστοφή, μαιευτήρα-χειρουργό και γυναικολόγο στο Αρεταίειο Νοσοκομείο για την συζήτηση που είχαμε σχετικά με την ασθένεια της προ-εκλαμψίας που εμφανίζουν οι έγκυες γυναίκες.

Τέλος θα ήθελα, να ευχαριστήσω τους γονείς μου, τους φίλους μου και τους συμφοιτητές μου για την αγάπη τους, την ανεκτίμητη υποστήριξη και συμπαράστασή τους αυτά τα πέντε ολόκληρα χρόνια των σπουδών μου.

Περίληψη

Η παρούσα ατομική διπλωματική εργασία ασχολήθηκε με το μοντέλο των Ασαφών γνωστικών Χαρτών – Fuzzy Cognitive Maps και τους αλγόριθμους για εκμάθηση του μοντέλου αυτού.

Στα πλαίσια της εργασίας έγινε κατανόηση του μοντέλου των Ασαφών γνωστικών Χαρτών, μελέτη διάφορων εφαρμογών/χρήσεων του και αλγορίθμων που χρησιμοποιούνται στην εκπαίδευση του.

Ακολούθως έγινε μελέτη της προ-εκλαμψίας, μιας σοβαρής επιπλοκής που συμβαίνει στις εγκύους γυναίκες. Η πρόβλεψη για εμφάνιση προ-εκλαμψίας σε οποιαδήποτε έγκυο γυναίκα με χρήση ενός μοντέλου Ασαφών γνωστικών χαρτών θα ήταν πολύ βοηθητική για τους γυναικολόγους – μαιευτήρες.

Στη συνέχεια έγινε υλοποίηση του αλγορίθμου Nonlinear Hebbian και τροποποίηση του γενικού γενετικού αλγορίθμου για εκπαίδευση του μοντέλου μας. Απώτερος στόχος ήταν το μοντέλο που θα προκύψει να μπορεί να κάνει πρόβλεψη της ασθένειας για περιπτώσεις εγκύων γυναικών.

Περιεχόμενα

Κεφάλαιο 1	Εισαγωγή.....	1
Κεφάλαιο 2	Ασαφείς Γνωστικοί Χάρτες	3
2.1	Ιστορικό Υπόβαθρο και Γενική Περιγραφή ΑΓΧ	3
2.2	Δομικά συστατικά ΑΓΧ	5
2.3	Ασαφής Λογική – Συναρτήσεις συμμετοχής	9
2.4	Στατική Ανάλυση ΑΓΧ	13
2.5	Δυναμική Ανάλυση του μοντέλου του ΑΓΧ	16
2.5.1	Συναρτήσεις Ανανέωσης Τιμών των Κόμβων	20
2.5.2	Συναρτήσεις Μεταφοράς	22
2.6	Διάφορες χρήσεις και εφαρμογές του μοντέλου	22
2.7	Αδυναμία Μοντέλου ΑΓΧ	23
2.8	Περιγραφή εργαλείου κατασκευής και εκτέλεσης Δυναμικής Ανάλυσης ΑΓΧ	24
2.9	Παράδειγμα για επεξήγηση και κατανόηση του μοντέλου του ΑΓΧ	28
Κεφάλαιο 3	Αλγόριθμοι Εκμάθησης ΑΓΧ	37
3.1	Γενικά - Αλγόριθμοι Εκμάθησης ΑΓΧ	37
3.2	Αλγόριθμος Εκμάθησης Nonlinear Hebbian Learning για ΑΓΧ	41
3.2.1	Περιγραφή Αλγορίθμου NHL	41
3.2.2	Καθορισμός Παραμέτρων Αλγορίθμου NHL	44
3.2.3	Συνθήκες Τερματισμού	45
3.2.4	Εφαρμογή αλγορίθμου NHL	47
3.3	Γενετικοί Αλγόριθμοι για εκμάθηση ΑΓΧ	49
3.3.1	Βασική Περιγραφή Γενετικών Αλγορίθμων	49
3.3.2	Μέθοδος Εκμάθησης μοντέλου ΑΓΧ με ΓΑ	50
3.3.3	Στοιχεία Περιβάλλοντος ΓΑ (τελεστές και άλλες παράμετροι)	52

Κεφάλαιο 4	Ανάλυση Δεδομένων	65
4.1	Τι είναι η Προ-εκλαμψία	65
4.1.1	Ενδείξεις και Συμπτώματα	66
4.1.2	Παράγοντες Κινδύνου	67
4.2	Δεδομένα	68
4.3	Παρουσίαση Ανάλυσης Δεδομένων	73
Κεφάλαιο 5	Υλοποίηση – Αποτελέσματα	86
5.1	Επεξεργασία και Κανονικοποίηση Δεδομένων	86
5.2	Προσέγγιση για εκπαίδευση του μοντέλου ΑΓΧ με τη χρήση γενετικών Αλγορίθμων	90
5.3	Προσέγγιση για εκπαίδευση του μοντέλου ΑΓΧ με τον NHL	95
5.4	Εκτέλεση Εκπαίδευσης και Αποτελέσματα	98
5.5	Επιπρόσθετες Δοκιμές Εκτέλεσης και Αποτελέσματα	101
Κεφάλαιο 6	Συμπεράσματα	114
Βιβλιογραφία		117
Π α ρ ά ρ τ η μ α Α		A-1
Π α ρ ά ρ τ η μ α Β		B-1

Κεφάλαιο 1

Εισαγωγή

Καθώς εξελίσσεται η επιστήμη της Πληροφορικής, όλο και περισσότερα υπολογιστικά συστήματα αναπτύσσονται για επίλυση προβλημάτων. Αν αναλογιστούμε ότι από την απλή υπολογιστική μηχανή η οποία εκτελούσε απλές αριθμητικές πράξεις, έχουμε φτάσει στο σημείο της ανάπτυξης συστημάτων που επιλύουν τα πολύπλοκα προβλήματα τα οποία είναι αδύνατο να λύσουν απλοί αλγόριθμοι ή μαθηματικά μοντέλα.

Τα συστήματα αυτά χαρακτηρίζονται ως ‘ευφυή’ μιας και μπορούν να προσαρμόζουν την ανθρώπινη, εξυπνάδα, γνώση, λογική, εμπειρία για την επίλυση ενός σύνθετου προβλήματος. Ένα τέτοιο σύστημα είναι και οι Ασαφείς Γνωστικοί Χάρτες. Οι Ασαφείς Γνωστικοί Χάρτες (ΑΓΧ) αποτελούν μια συστηματική διαδικασία συλλογής, κωδικοποίησης και εξέλιξης πληροφοριών σε σχέση με διάφορα φαινόμενα στη περιοχή μοντελοποίησης και την αλληλεπίδραση τους, τόσο εσωτερική όσο και εξωτερική με το περιβάλλον. Μοντελοποιώντας λοιπόν ένα πολύπλοκο σύστημα με ένα ΑΓΧ μπορεί να δώσει στο χρήστη αποτελέσματα δυναμικής πρόγνωσης καταστάσεων και ρύθμισης του μοντελοποιημένου προβλήματος. Το μοντέλο μπορεί να παράγει συμπεράσματα μέσα σε συνθήκες αβεβαιότητας, όπως κάνουν οι εμπειρογνώμονες.

Για την δημιουργία των Ασαφών γνωστικών Χαρτών, απαραίτητο στοιχείο είναι η γνώση. Οι έμπειροι και οι ειδικοί, του συγκεκριμένου πεδίου του οποίου γίνεται η ανάπτυξη συστήματος, είναι αυτοί που δίνουν την γνώση για τη δημιουργία του.

Το δυσκολότερο και σημαντικότερο κομμάτι της δημιουργίας των Ασαφών γνωστικών Χαρτών είναι ο σωστός καθορισμός των παραμέτρων (μεταβλητών) που περιγράφουν το σύστημα από τους έμπειρους. Δύσκολος είναι και ο καθορισμός των σχέσεων που συσχετίζουν τις παραμέτρους μεταξύ τους και γίνεται και αυτό από τους ειδικούς.

Αυτή είναι και η κύρια αδυναμία του μοντέλου των Ασαφών Γνωστικών Χαρτών αφού πολλές φορές δεν μπορεί να γίνει σωστός καθορισμός των παραμέτρων και των σχέσεων από τους ειδικούς.

Τα τεχνητά νευρωνικά δίκτυα, τα οποία αποτελούν και αυτά ένα ευφυή υπολογιστικό σύστημα και τα οποία χρησιμοποιούνται στην επίλυση πολύπλοκων προβλημάτων είναι ικανά να "μαθαίνουν" και να αποκρίνονται με "εξυπνάδα" σε νέα ερεθίσματα. Για να βελτιώσουν την αδυναμία της δημιουργίας των Ασαφών Γνωστικών Χαρτών προτάθηκαν αλγόριθμοι για πιο αυτοματοποιημένη δημιουργία του. Μερικοί αλγόριθμοι που έχουν προταθεί χρησιμοποιούν κανόνες εκμάθησης των τεχνητών νευρωνικών δικτύων και εκτελούν μη επιβλεπόμενη μάθηση για την δημιουργία του μοντέλου. Άλλοι αλγόριθμοι βασίζονται στην εξελισσόμενη υπολογιστική και τους γενετικούς αλγόριθμους. Εδώ τίθεται το ερώτημα κατά πόσο οι αλγόριθμοι αυτοί μπορούν να εκπαιδεύσουν το μοντέλο των Ασαφών Γνωστικών χαρτών καθορίζοντας τις σχέσεις με τέτοιο τρόπο ώστε να κάνει πάντα σωστή πρόβλεψη.

Τα μοντέλα των Ασαφών Γνωστικών Χαρτών χρησιμοποιούνται σε πλήθος εφαρμογών, από την εξαγωγή συμπερασμάτων σε πολιτικό-οικονομικούς παράγοντες μιας χώρας, μέχρι την διάγνωση μιας ιατρικής διαταραχής.

Η έγκαιρη πρόβλεψη της προ-εκλαμψίας, μιας πολύ σοβαρής επιπλοκής που συμβαίνει μόνο στις εγκύους γυναίκες αποτελεί ένα πραγματικό πρόβλημα. Θα ήταν μεγάλη βοήθεια για τους γυναικολόγους, γιατρούς αλλά και τις έγκυες γυναίκες η ανάπτυξη ενός υπολογιστικού συστήματος που θα μπορεί να προβλέψει αν μια έγκυος γυναίκα θα εμφανίσει ή όχι προ-εκλαμψία.

Μέσα από τη μελέτη αυτή γίνεται μια προσπάθεια για ανάπτυξη ενός μοντέλου Ασαφή γνωστικού Χάρτη, με εφαρμογή αλγορίθμων εκμάθησης και το οποίο θα είναι επιτυχημένο εάν μπορεί να προβλέψει σωστά τις περιπτώσεις των εγκύων γυναικών που θα εμφανίσουν προ-εκλαμψία.

Κεφάλαιο 2

Ασαφείς Γνωστικοί Χάρτες

2.1 Ιστορικό Υπόβαθρο και Γενική Περιγραφή ΑΓΧ	3
2.2 Δομικά συστατικά ΑΓΧ	5
2.3 Ασαφής Λογική – Συναρτήσεις συμμετοχής	9
2.4 Στατική Ανάλυση ΑΓΧ	13
2.5 Δυναμική Ανάλυση του μοντέλου του ΑΓΧ	16
2.5.1 Συναρτήσεις Ανανέωσης Τιμών των Κόμβων	20
2.5.2 Συναρτήσεις Μεταφοράς	22
2.6 Διάφορες χρήσεις και εφαρμογές του μοντέλου	22
2.7 Αδυναμία Μοντέλου ΑΓΧ	23
2.8 Περιγραφή εργαλείου κατασκευής και εκτέλεσης δυναμικής ανάλυσης ΑΓΧ	24
2.9 Παράδειγμα για επεξήγηση και κατανόηση του μοντέλου του ΑΓΧ	28

Στο κεφάλαιο αυτό θα αναλύσουμε τις παραμέτρους που χαρακτηρίζουν το μοντέλο των ΑΓΧ, τις μεταξύ τους αλληλεπιδράσεις και άλλες ιδιότητες του μοντέλου. Ακόμα θα γίνει αναφορά σε απλό παράδειγμα για καλύτερη κατανόηση του μοντέλου και θα αναφερθεί το εργαλείο για την προσομοίωση του μοντέλου.

2.1 Ιστορικό Υπόβαθρο και Γενική Περιγραφή ΑΓΧ

Οι ασαφείς γνωστικοί χάρτες είναι μια ευέλικτη υπολογιστική μέθοδος η οποία χρησιμοποιείται για να μοντελοποιήσει πολύπλοκα συστήματα και η οποία συνδυάζει στοιχεία από τους χώρους της Ασαφούς Λογικής και των Νευρωνικών Δικτύων. Είναι σχετικά μια πρόσφατη μέθοδος μιας και η εμφάνιση της μπορεί να καταγραφεί στις υπολογιστικές επιτεύξεις του 19^{ου} αιώνα [6,10,20].

Οι ασαφείς γνωστικοί χάρτες αποτελούν μια επέκταση των γνωστικών χαρτών. Οι γνωστικοί χάρτες προτάθηκαν από τον πολιτικό επιστήμονα Robert Axelrod το 1970 για αναπαράσταση κοινωνικής και επιστημονικής γνώσης και για να περιγράψει μεθόδους που χρησιμοποιήθηκαν στην λήψη αποφάσεων σε κοινωνικά και πολιτικά συστήματα. Οι γνωστικοί χάρτες του Axelrod ήταν γράφοι που αποτελούνταν από κόμβους και ακμές με θετικό ή αρνητικό πρόσημο. Οι κόμβοι αποτελούσαν μεταβλητές έννοιες όπως για παράδειγμα κοινωνική αστάθεια ενώ οι ακμές αποτελούσαν σχέσεις αιτίας –αποτελέσματος μεταξύ των κόμβων. Μια ακμή με θετικό πρόσημο αποτελεί μια θετική σχέση αιτιότητας από τον κόμβο Α στον κόμβο Β και υπονοεί ότι αν αυξηθεί το Α αυξάνεται και το Β, ενώ αν μειωθεί το Α μειώνεται και το Β. Αντίθετα υπάρχει μια αντίστροφη σχέση όταν η ακμή έχει αρνητικό πρόσημο και άρα αποτελεί αρνητική σχέση αιτιότητας, δηλαδή όταν αυξηθεί το Α μειώνεται το Β ενώ αν μειωθεί το Α αυξάνεται το Β. [1]

Στη συνέχεια ο Bart Kosko, συγγραφέας, ηλεκτρολόγος μηχανικός στο Πανεπιστήμιο της Νότιας Καλιφόρνιας και γνωστός ερευνητής σε ασαφή λογική και νευρωνικά δίκτυα εμπλούτισε τους γνωστικούς χάρτες στους ασαφείς γνωστικούς χάρτες (1986). Συγκεκριμένα οι ασαφείς γνωστικοί χάρτες του Kosko έχουν ασαφείς τιμές στους κόμβους και ασαφείς βαθμούς στις διασυνδέσεις μεταξύ των κόμβων [1].

Οι ΑΓΧ χρησιμοποιούν συμβολική αναπαράσταση για την περιγραφή και τη μοντελοποίηση των πολύπλοκων συστημάτων. Συγκεκριμένα η γραφική τους αναπαράσταση είναι ένας προσημασμένος κατευθυνόμενος γράφος, ο οποίος αποτελείται από κόμβους και ακμές που συνδέουν τους κόμβους αυτούς και οι οποίες έχουν κάποιο βάρος. Οι κόμβοι αναπαριστούν χαρακτηριστικά συμπεριφοράς του συστήματος ενώ τα βάρη τα οποία έχουν κατεύθυνση, πρόσημο και τιμή αναπαριστούν μια σχέση αιτιότητας που υπάρχει μεταξύ των κόμβων [20].

Η μοντελοποίηση ενός συστήματος σε έναν ΑΓΧ μπορεί να χαρακτηριστεί ως ένα μονο-επίπεδο ή πολύ-επίπεδο δίκτυο όπου οι κόμβοι αντιπροσωπεύουν βασικές έννοιες του συστήματος και τα βάρη συμβολίζουν κάποια σχέση αιτίας-συνέπειας ανάμεσα σε αυτές τις έννοιες. Ο ΑΓΧ σαν μοντέλο μπορεί να δώσει στο χρήστη μεγάλη δυναμική πρόγνωσης καταστάσεων ή ρυθμίσεων του μοντελοποιημένου συστήματος, αφού

επιτρέπει την ανάδραση ανάμεσα στο χρόνο και άρα τη δυνατότητα μοντελοποίησης συστημάτων που εξελίσσονται με το χρόνο. Η ανάδραση αυτή παίζει ένα πολύ σημαντικό ρόλο στην ανανέωση των τιμών των καταστάσεων των κόμβων διαδίδοντας αιτιώδεις επιδράσεις από κόμβο σε κόμβο χρησιμοποιώντας ποικίλα μονοπάτια [20].

Η ανάπτυξη των μοντέλων των ασαφών γνωστικών χαρτών βασίζεται κυρίως στην γνώση και εμπειρία κάποιων ειδικών που είναι σχετικοί με τον τομέα του συστήματος που μοντελοποιείται και οι οποίοι γνωρίζουν την συμπεριφορά του συστήματος σε διάφορες περιστάσεις. Ο καθορισμός των κόμβων καθώς και των διασυνδέσεων μεταξύ αυτών των κόμβων εναπόκειται στην γνώση και την εμπειρία των ειδικών. Για μια επιτυχημένη κατασκευή ενός ΑΓΧ θεωρούνται σημαντικές παράμετροι ο αριθμός των κόμβων, η πυκνότητα του δικτύου και ο σωστός καθορισμός των τιμών ενεργοποίησης των κόμβων αλλά και των τιμών των βαρών [10, 20].

2.2 Δομικά συστατικά ΑΓΧ

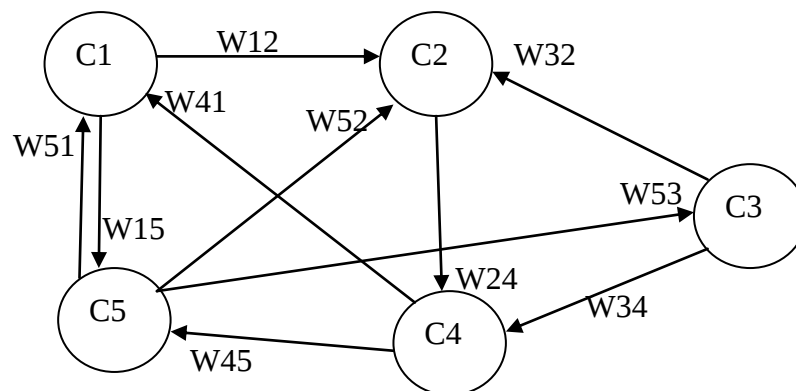
Οι κόμβοι αποτελούν ένα από τα δομικά συστατικά του ασαφή γνωστικού χάρτη και αναπαριστούν διαφορετικές έννοιες που διέπουν το σύστημα που μοντελοποιείται. Οι έννοιες αυτές ή αλλιώς χαρακτηριστικά του συστήματος μπορεί να είναι μεταβλητές, γεγονότα, καταστάσεις, παράγοντες, πράξεις, στόχοι, τάσεις, περιορισμοί του συστήματος κτλ. [22] Οι κόμβοι χαρακτηρίζονται από μια αριθμητική τιμή, την τιμή καταστάσεως ενεργοποίησης A_i . Η τιμή αυτή αντιπροσωπεύει μια ποσοτική μέτρηση της «παρουσίας-υπόστασης» της έννοιας/χαρακτηριστικού στο μοντέλο. Άρα ένας κόμβος με μια ψηλή αριθμητική τιμή σημαίνει ότι η έννοια/χαρακτηριστικό που αναπαριστά έχει ισχυρή-έντονη παρουσία στο σύστημα. [20]

Για παράδειγμα αν η αριθμητική τιμή ενός κόμβου είναι θετική και πολύ ψηλή (0.9) τότε σημαίνει ότι η κατάσταση του αντίστοιχου χαρακτηριστικού (που αναπαριστά ο κόμβος) έχει αυξηθεί πολύ ή βρίσκεται σε πολύ ψηλά μη φυσιολογικά θετικά επίπεδα. Εάν η αριθμητική τιμή είναι θετική αλλά χαμηλή (0.2) τότε η αύξηση στη κατάσταση του χαρακτηριστικού είναι σχετικά μικρή. Παρόμοια αν η αριθμητική τιμή του κόμβου είναι αρνητική και πολύ ψηλή (-0.9) σημαίνει ότι η κατάσταση του χαρακτηριστικού έχει μειωθεί σημαντικά ή φθίνει σε πολύ χαμηλά αρνητικά επίπεδα. Ενώ αν η τιμή

είναι αρνητική και χαμηλή (-0.2) σημαίνει ότι η μείωση στη κατάσταση του αντίστοιχου χαρακτηριστικού είναι σχετικά μικρή.[6]

Κατά την ανάπτυξη του μοντέλου οι ειδικοί σύμφωνα με την γνώση που έχουν γύρω από την περιοχή της εφαρμογής αναγνωρίζουν τις βασικές έννοιες, και τα χαρακτηριστικά του συστήματος και δίνουν μια γλωσσολογική περιγραφή για την τιμή ενεργοποίησης στο κάθε ένα από αυτά. Οι αριθμητικές τιμές καταστάσεως ενεργοποίησης αποτελούν αποτέλεσμα μετασχηματισμού των πραγματικών τιμών των μεταβλητών (έννοιών/ χαρακτηριστικών) του συστήματος στο διάστημα $[0,1]$. Στο παράδειγμα που αναφέρεται πιο πάνω οι τιμές καταστάσεων των κόμβων παίρνουν τιμές στο διάστημα $[-1,1]$. Αυτό εξαρτάτε από το σύστημα που μοντελοποιείται και από το τι θα επιλέξουν οι ειδικοί. Ο μετασχηματισμός αυτός γίνεται για να αποφευχθεί η ανάθεση δυσανάλογης σημαντικότητας σε συγκεκριμένους κόμβους, αφού η ανάθεση της τιμής είναι υποκειμενική και εναπόκειται στη γνώμη του κάθε ειδικού ξεχωριστά. Έτσι ενδέχεται να υπάρξει μεγάλη απόκλιση στις τιμές ενεργοποίησης κάθε κόμβου και αυτό έχει ως αποτέλεσμα να υποβαθμιστεί η παρουσία ενός κόμβου στο γενικό σύνολο του συστήματος. Για τον μετασχηματισμό των τιμών καταστάσεων ενεργοποίησης στο διάστημα $[0,1]$ χρησιμοποιούνται συναρτήσεις μεταφοράς τις οποίες θα αναφέρουμε στη συνέχεια. [10, 20]

Ο μαθηματικός τρόπος αναπαράστασης των τιμών καταστάσεων ενεργοποίησης είναι σε μορφή διανύσματος. Αν δηλαδή είχαμε ένα ΑΓΧ που μοντελοποιούσε ένα σύστημα με n κόμβους (μεταβλητές) οι τιμές καταστάσεων ενεργοποίησης των κόμβων αυτών θα ήταν ένα $1 \times n$ διάνυσμα A . Στο παράδειγμα του σχήματος 2.2.1 το διάνυσμα τιμών καταστάσεων ενεργοποίησης θα ήταν ως εξής: $A = [A_1, A_2, A_3, A_4, A_5]$. [20]

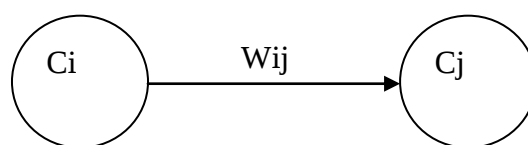


Σχήμα 2.2.1 Παράδειγμα μοντέλου ΑΓΧ

Στο μοντέλο του ασαφή γνωστικού χάρτη, οι κόμβοι είναι διασυνδεδεμένοι μεταξύ τους με ακμές. Συγκεκριμένα μια ακμή συνδέει δύο κόμβους μεταξύ τους και χαρακτηρίζεται από κατεύθυνση, πρόσημο και τιμή(βάρος). Οι ακμές είναι το δεύτερο συστατικό των ΑΓΧ. [20]

Κάθε μια ακμή που συνδέει δύο κόμβους περιγράφει μια σχέση αιτίας συνέπειας (αποτελέσματος) που υφίσταται μεταξύ των δύο κόμβων, δηλαδή το αποτέλεσμα που έχει η μεταβολή της τιμής ενεργοποίησης καταστάσεως του ενός κόμβου στη διαμόρφωση της τιμής ενεργοποίησης καταστάσεως του διασυνδεδεμένου κόμβου. Η κατεύθυνση της ακμής υπονοεί ποιος κόμβος επηρεάζει ποιόν, δηλαδή ποιος κόμβος θεωρείται η αιτία και ποιος η συνέπεια (αποτέλεσμα). Το πρόσημο της σχέσης αιτιότητας μπορεί να είναι είτε θετικό είτε αρνητικό. Εάν είναι θετικό τότε σημαίνει ότι υπάρχει ευθεία σχέση επηρεασμού ανάμεσα στους δύο κόμβους ενώ αν το πρόσημο είναι αρνητικό τότε σημαίνει ότι υπάρχει αντίστροφη σχέση επηρεασμού ανάμεσα στους δύο κόμβους. Τέλος υπάρχει το βάρος που χαρακτηρίζει τις σχέσεις αυτές και στην ουσία συμβολίζει το βαθμό που επηρεάζει ο ένας κόμβος τον άλλο. Μεταξύ των κόμβων υπάρχουν τρεις πιθανοί τύποι σχέσης αιτιότητας που εκφράζουν τον τρόπο που επηρεάζει ο ένας κόμβος τον άλλο. [10, 20]

Έστω έχουμε δύο κόμβους C_i και C_j και συνδέονται με μια ακμή που έχει κατεύθυνση από τον C_i στον C_j και βάρος W_{ij} , (σχήμα 2.2.2)



Σχήμα 2.2.2 Παράδειγμα διασύνδεσης δύο κόμβων

Οι τρεις περιπτώσεις για το βάρος W_{ij} που δείχνουν τον τρόπο που επηρεάζει ο κόμβος C_i τον κόμβο C_j είναι οι ακόλουθες:

1. Αν το $W_{ij} > 0$ τότε,

Εάν η τιμή της καταστάσεως ενεργοποίησης του κόμβου C_i , A_i αυξηθεί τότε θα αυξηθεί και η τιμή της καταστάσεων του κόμβου C_j , A_j . Αν η τιμή της

καταστάσεως ενεργοποίησης του κόμβου C_i , A_i , μειωθεί τότε θα μειωθεί και η τιμή της καταστάσεως ενεργοποίησης του κόμβου C_j , A_j .

2. Αν το $W_{ij} < 0$ τότε,

Εάν η τιμή της καταστάσεως ενεργοποίησης του κόμβου C_i , A_i αυξηθεί τότε θα μειωθεί η τιμή της καταστάσεων του κόμβου C_j , A_j . Αν η τιμή της καταστάσεως ενεργοποίησης του κόμβου C_i , A_i , μειωθεί τότε θα αυξηθεί η τιμή της καταστάσεως ενεργοποίησης του κόμβου C_j , A_j .

3. Αν το $W_{ij} = 0$ τότε,

Δεν υπάρχει σχέση αιτιότητας μεταξύ του κόμβου C_i και του κόμβου C_j . [20]

Κατά την ανάπτυξη και κατασκευή του μοντέλου του ΑΓΧ οι ειδικοί εκτός από τον καθορισμό των κύριων χαρακτηριστικών του συστήματος που αναπαριστούν οι κόμβοι και τις αρχικές τιμές καταστάσεως ενεργοποίησης τους, καθορίζουν τις σχέσεις αιτιότητας μεταξύ των κόμβων αυτών, το είδος της σχέσης, αν θα είναι θετική ή αρνητική και τον βαθμό της σχέσης, πόσο ισχυρή είναι. Λόγω του ότι οι σχέσεις αιτιότητας καθορίζονται από τους ειδικούς οι οποίοι τις περιγράφουν με τη χρήση λεκτικών μεταβλητών χαρακτηρίζονται ασαφείς και ανακριβείς. Στη συνέχεια οι σχέσεις αυτές ασαφοποιούνται με τη χρήση συναρτήσεων συμμετοχής, και οι λεκτικές μεταβλητές μετατρέπονται σε αριθμητικές μεταβλητές παίρνοντας τιμές στο πραγματικό διάστημα $[-1,1]$. Έτσι η αναπαράσταση της αιτιότητας επιτρέπει την διαβάθμιση της σε διάφορα επίπεδα, αφού τα βάρη μπορούν να πάρουν τιμές στο διάστημα $[-1,1]$ και αυτό κάνει το μοντέλο να διαφέρει από τους γνωστικούς χάρτες που πρότεινε ο Axelrod στους οποίους επικρατούσε η δυαδική λογική 0, 1 στις σχέσεις αιτιότητας. Η αριθμητική τιμή του βάρους της σχέσης εκφράζει το βαθμό – ένταση της αιτιώδους επιρροής της συγκεκριμένης σχέσης, δηλαδή πόσο ισχυρή είναι η σχέση μεταξύ δύο κόμβων και πόση επιρροή έχει ο ένας κόμβος στον άλλο. Η αριθμητική τιμή -1 έχει ολοκληρωτικά αρνητική επιρροή, ενώ η αριθμητική τιμή 1 δηλώνει ολοκληρωτικά θετική επιρροή. Όταν η τιμή είναι 0 σημαίνει ότι δεν υπάρχει σχέση αιτιότητας ενώ οι υπόλοιπες ενδιάμεσες τιμές αντιστοιχούν σε διαφορετικά επίπεδα αιτιότητας. Όσο η αριθμητική τιμή του βάρους είναι πιο κοντά στο 0 τόσο πιο αδύνατη είναι η σχέση, όσο πιο κοντά είναι στο 1 τόσο πιο έντονα επηρεάζει ευθέως ο ένας κόμβος τον άλλο, ενώ όσο πιο κοντά είναι στο -1 τόσο πιο έντονα επηρεάζει αντιστρόφως ο ένας κόμβος τον άλλο. [10, 20]

Ο μαθηματικός τρόπος συμβολισμού των βαρών του μοντέλου ενός ΑΓΧ γίνεται με ένα τετράγωνο δυσδιάστατο πίνακα. Αν για παράδειγμα το μοντέλο του ΑΓΧ μοντελοποιεί ένα σύστημα με n μεταβλητές, οι τιμές των βαρών είναι ένας $n \times n$ πίνακας W . Η θέση W_{ij} υπονοεί ότι υπάρχει μια σχέση από τον κόμβο C_i προς τον C_j με αριθμητική τιμή e_{ij} .

Στο παράδειγμα του σχήματος 2.2.1 ο πίνακας των βαρών του μοντέλου θα ήταν ως ακολούθως:

	C1	C2	C3	C4	C5
C1	0	W12	0	0	W15
C2	0	0	0	W24	0
C3	0	W32	0	W34	0
C4	W41	0	0	0	W45
C5	W51	0	W53	0	0

Πίνακας 2.2.3 Πίνακας βαρών παραδείγματος σχήματος 2.2.1

Παρατήρηση:

Οι ακμές μεταξύ των κόμβων συμβολίζουν σχέσεις αιτίας αποτελέσματος, θεωρείται ότι δεν μπορεί να υπάρξει σχέση αιτιότητας μεταξύ ενός κόμβου και του εαυτού του και άρα δεν υφίσταται ακμή μεταξύ ενός κόμβου και του εαυτού του οπότε και η θέση W_{ii} στον πίνακα θα έχει τιμή ίση με μηδέν.

2.3 Ασαφής Λογική – Συναρτήσεις συμμετοχής

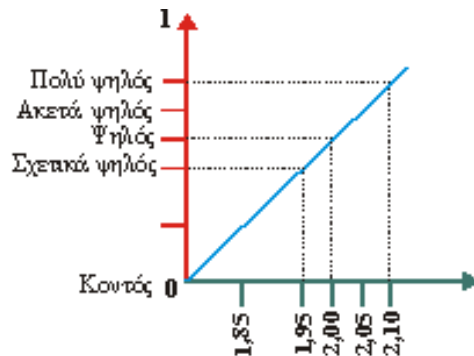
Οι ασαφείς γνωστικοί χάρτες εκτός από το χώρο των Νευρωνικών Δικτύων παίρνουν στοιχεία από το χώρο της Ασαφούς Λογικής.

Η ασαφής Λογική η οποία πρωτοπαρουσιάστηκε από τον Lofti Zadeh το 1965 μάς παρέχει ένα σύστημα κανόνων για να αντιμετωπίσουμε προβλήματα στα οποία η κλασική Λογική κρίνεται ανεπαρκής. [38]

Ένα παράδειγμα προβλήματος το οποίο δε θα μπορούσε να λύσει η κλασική λογική είναι το εξής: ας υποθέσουμε ότι θέλουμε να ταξινομήσουμε τους παίκτες της εθνικής

ομάδας μπάσκετ με κριτήριο το ύψος τους. Ένας παίκτης που έχει ύψος 1,95 θα χαρακτηριστεί σχετικά ψηλός ενώ ένας που έχει ύψος 2,05, αρκετά ψηλός κλπ.

Υπάρχουν βέβαια κι άλλες διαβαθμίσεις: ένας παίκτης με ύψος 2,15 θα μπορούσε να χαρακτηριστεί παρά πολύ ψηλός, κι ένας με ύψος 2,20, πανύψηλος.[38] (σχήμα 2.3.1)



Σχήμα 2.3.1 Γραφική αναπαράσταση διαβάθμισης της μεταβλητής του ύψους

Οι λέξεις (λεκτικές μεταβλητές) ή οι φράσεις-κλειδιά: σχετικά, αρκετά, πάρα πολύ, πανύψηλος δεν μπορούν να αποδοθούν με την κλασική Λογική. Ο λόγος είναι ότι η κλασική λογική βασίζεται μόνο σε δύο τιμές: το ναι και το όχι. Αυτό σημαίνει ότι ένας μπάσκετμπολίστας θα είναι ή ψηλός ή κοντός και δηλαδή ή θα είναι μέλος του συνόλου των ψηλών ή του συνόλου των κοντών δεν υπάρχει ενδιάμεση κατάσταση. [38]

Όμως, συχνά στην καθημερινή μας ζωή όταν περιγράφουμε κάτι χρησιμοποιούμε όρους που κινούνται ανάμεσα σ' αυτά τα δύο άκρα. Κάτι ανάλογο συμβαίνει και με τα χρώματα: ανάμεσα στο άσπρο και το μαύρο υπάρχουν πολλοί τόνοι του γκριζου. [38] (σχήμα 2.3.2).



Σχήμα 2.3.2 Διαβαθμίσεις του γκριζου χρώματος

Η ασαφής Λογική (Fuzzy Logic) είναι μια λογική των διαβαθμίσεων. Εξετάζει σε ποιο βαθμό ένα στοιχείο αποτελεί μέλος ενός συνόλου. Έτσι, έχουμε τη δυνατότητα να αντιστοιχίσουμε στους όρους που είδαμε προηγουμένως (σχετικά ψηλός, πολύ ψηλός

κτλ) σε αριθμητικές τιμές που κυμαίνονται ανάμεσα στο 0 (κοντός) και το 1 (πανύψηλος). [38]

Για παράδειγμα, ένας παίκτης με ύψος 2,10 παίρνει την τιμή 0,8. Η τιμή αυτή δείχνει τον βαθμό συμμετοχής του στο σύνολο ψηλός.

Γενικά η αντιστοίχιση των λεκτικών μεταβλητών σε αριθμητικές τιμές γίνεται με τη χρήση των συναρτήσεων συμμετοχής. [38]

Συνάρτηση συμμετοχής $\mu_A(x)$ ορίζεται ως ο βαθμός στον οποίο το αντικείμενο x ενός συνόλου X ανήκει στο σύνολο A , έτσι ώστε $\mu_A(x) = \text{Degree}(x \in A)$. Κάθε συνάρτηση συμμετοχής αναπαριστάται γραφικά ως μια καμπύλη και η οποία αντιστοιχεί ένα στοιχείο από το χώρο εισόδου με μία τιμή βαθμού συμμετοχής στο πεδίο 0 και 1. Μια συνάρτηση συμμετοχής μπορεί να οριστεί με τέτοιο τρόπο ώστε να ικανοποιεί το ασαφές σύνολο που ορίζεται μέσα από το σύστημα που μοντελοποιείται σε κάθε περίπτωση. [20]

Ένα ασαφές σύνολο είναι μία προέκταση του κλασσικού συνόλου. Έστω X ένα τυπικό σύνολο αντικειμένων που το καλούμε σύμπαν αναφοράς και του οποίου τα στοιχεία τα συμβολίζουμε με x . Τότε το ένα ασαφές σύνολο A είναι ένα υποσύνολο στο σύνολο X και το οποίο δεν έχει σαφή όρια και αποτελείται από διατεταγμένα ζευγάρια ως εξής:

$$A = \{x, \mu_A(x) \mid x \in X\},$$

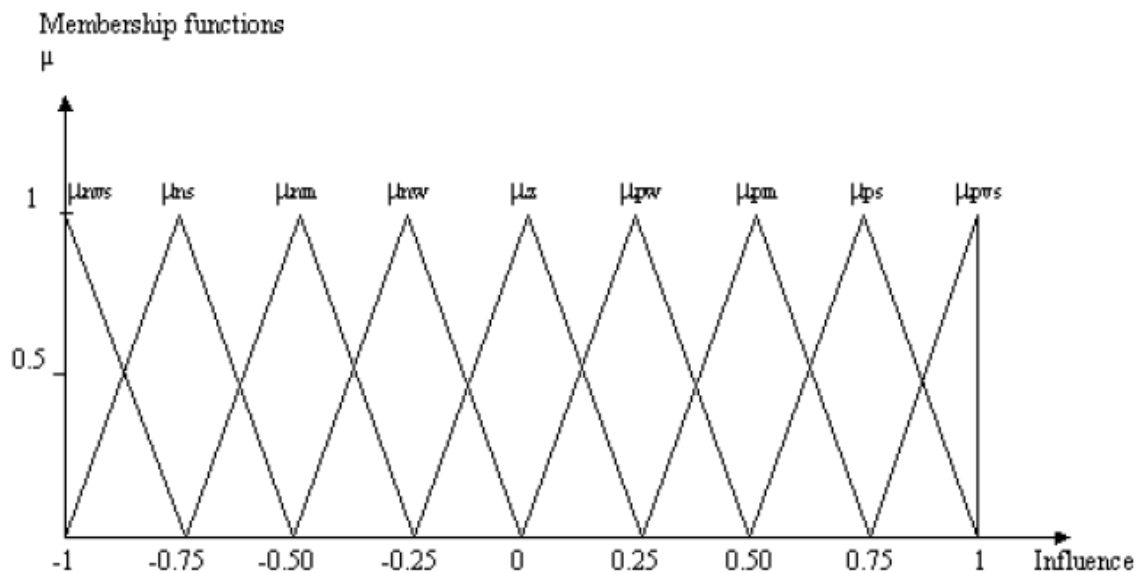
Όπου το $\mu_A(x)$ είναι η συνάρτηση συμμετοχής του στοιχείου x στο σύνολο A , δηλαδή σε ποιο βαθμό/ποσοστό ανήκει το x στο A . Ένα στοιχείο x είναι δυνατόν να ανήκει παράλληλα και σε άλλο ασαφές σύνολο B αλλά με διαφορετικό βαθμό συμμετοχής $\mu_B(x)$. [20]

Στο παράδειγμα με τους μπασκετμπολίστες η συνάρτηση συμμετοχής θα μπορούσε να ήταν μια καμπύλη όπου ο άξονας y περιλαμβάνει τις τιμές από 0 μέχρι το 1 και ορίζει μια ομαλή μετάβαση από τον πολύ κοντό στον πανύψηλο παίκτη. Η ιδέα της χρήσης

των συναρτήσεων συμμετοχής θεωρεί ότι δύο μπασκετμπολίστες διαφορετικού ύψους είναι ψηλοί αλλά σε διαφορετικό βαθμό ο ένας από τον άλλο.

Στην περίπτωση του μοντέλου των AGX, όπου η ανάπτυξη και η κατασκευή του δικτύου γίνεται από τους ειδικούς οι συναρτήσεις συμμετοχής χρειάζονται για την μετατροπή της γλωσσολογικής περιγραφής σε αριθμητική τιμή. Η ομάδα των ειδικών καθορίζει τον αριθμό και το είδος των κόμβων που περιγράφουν τον AGX. Οι ειδικοί μέσα από την εμπειρία τους γνωρίζουν ποιοι είναι οι κύριοι παράγοντες που περιγράφουν την συμπεριφορά του συστήματος και κάθε ένα από αυτούς τους παράγοντες παριστάνεται σε έναν κόμβο στο μοντέλο του AGX. Επίσης γνωρίζουν ποια στοιχεία/μεταβλητές του συστήματος επηρεάζουν άλλα στοιχεία μεταβλητές και για τους αντίστοιχους κόμβους καθορίζουν την αρνητική ή θετική επιρροή από τον ένα κόμβο στον άλλο με έναν ασαφή βαθμό διασύνδεσης-σχέσης αιτιότητας. [10, 20]

Η μεθοδολογία ανάπτυξης του AGX βασίζεται στις ασαφείς εκφράσεις για την περιγραφή των αλληλεξαρτήσεων μεταξύ των κόμβων. Οι ειδικοί όταν περιγράφουν τις σχέσεις αιτιότητας καθορίζουν αρχικά την επιρροή του ενός κόμβου σε άλλον ως αρνητική ή θετική και στη συνέχεια αξιολογούν έπειτα το βαθμό στον οποίο επηρεάζει ο ένας κόμβος τον άλλο με τη χρήση μιας λεκτικής μεταβλητής όπως, «δυνατή επίδραση», «μεσαία επίδραση», «ασθενής επίδραση» κλπ. Συγκεκριμένα οι τιμές των βαρών καθορίζονται χρησιμοποιώντας τη μεταβλητή Επιρροή (Influence) η οποία παίρνει τιμές στο διάστημα $U = [-1, 1]$. Ο όρος T(Επιρροή) προτείνεται για να περιλάβει εννέα λεκτικές μεταβλητές. Με αυτές τις εννέα λεκτικές μεταβλητές ο ειδικός μπορεί να περιγράψει λεπτομερώς την επιρροή ενός κόμβου σε έναν άλλο και μπορεί να διακρίνει μεταξύ των διαφορετικών βαθμών επιρροής. Το σύνολο των εννέα λεκτικών μεταβλητών είναι το εξής: T(επιρροή) = {αρνητικά πολύ ισχυρό, αρνητικά ισχυρό, αρνητικά μέσο, αρνητικά αδύνατο, μηδέν, θετικά αδύνατο, θετικά μέσο, θετικά ισχυρό, θετικά πολύ ισχυρό}. Οι αντίστοιχες συναρτήσεις συμμετοχής για αυτές τις λεκτικές μεταβλητές παρουσιάζονται στο σχήμα 2.3.3. Και είναι οι ακόλουθες: μ_{nvs} , μ_{ns} , μ_{nm} , μ_{nw} , μ_z , μ_{pw} , μ_{pm} , μ_{ps} και μ_{pvs} . [8,10, 20, 22]



Σχήμα 2.3.3 Συναρτήσεις Συμμετοχής της λεκτικής μεταβλητής Influence(Επιρροή) [8]

Οι ειδικοί επιλέγουν μια ασαφή λεκτική μεταβλητή από το σύνολο T , για να περιγράψουν την σχέση μεταξύ δύο κόμβων. Στη συνέχεια οι λεκτικές μεταβλητές που προτείνονται από τους ειδικούς αθροίζονται χρησιμοποιώντας την μέθοδο SUM (Lin and Lee, 1996) και παράγεται ένα γενικό λεκτικό βάρος, το οποίο με μια μέθοδο αποασαφοποίησης όπως είναι η αποασαφοποίηση κέντρου βάρους (CoG) ή η αποασαφοποίηση κέντρου περιοχής (CoA) μετασχηματίζεται σε ένα αριθμητικό βάρος που ανήκει στο διάστημα $[-1, 1]$. [8,10, 20, 22]

2.4 Στατική Ανάλυση ΑΓΧ

Η στατική ανάλυση ενός ΑΓΧ περιλαμβάνει κάποια χαρακτηριστικά τα οποία δίνουν πληροφορίες για το μοντέλο του Ασαφής Γνωστικού Χάρτη. Συγκεκριμένα οι υπολογισμοί των χαρακτηριστικών αυτών που θα αναφερθούν στη συνέχεια μπορούν να γίνουν αφού κατασκευαστεί το μοντέλο από τους ειδικούς σαν έλεγχος για την ορθότητα της δομής του μοντέλου. [20]

Αυτά τα χαρακτηριστικά είναι τα εξής:

1. Πυκνότητα: Η πυκνότητα είναι ένα μέτρο που καθορίζει το βαθμό συνδετικότητας του δικτύου του ΑΓΧ, είναι το ποσοστό των υπαρχών

συνδέσεων μεταξύ των κόμβων προς όλες τις πιθανές συνδέσεις που μπορεί να υπάρξουν δεδομένου του αριθμού των κόμβων. Για τον υπολογισμό της χρειαζόμαστε τον αριθμό των κόμβων (N) και τον αριθμό των σχέσεων αιτιότητας (R) μεταξύ των κόμβων που έχει το μοντέλο του ΑΓΧ. Ο τύπος που την υπολογίζει είναι :

$$D = \frac{R}{N^2} \quad (1),$$

Η πυκνότητα δείχνει το βαθμό πολυπλοκότητας του δικτύου. Όσο πιο μεγάλη τιμή έχει τόσο πιο πολύπλοκο είναι το δίκτυο μας και το αντίστοιχο πρόβλημα που μοντελοποιεί. Έτσι η συνιστώμενη τιμή πυκνότητας για να έχει ένα δίκτυο πρέπει να ανήκει στο διάστημα [0.05, 0.3]. [20]

2. Σχετική «σημαντικότητα» κάθε κόμβου του δικτύου: Η κεντρικότητα όπως περιγράφεται από τον Kosko [1] είναι ένα μέτρο που καθορίζει πόσο σημαντικός είναι ένας κόμβος στο μοντέλο του ΑΓΧ. Ο υπολογισμός της κεντρικότητας ενός κόμβου C_j στο μοντέλο του ΑΓΧ είναι το άθροισμα των βαρών των εισερχόμενων ακμών σε αυτόν τον κόμβο (οι ακμές που δείχνουν προς τον κόμβο C_j) και το άθροισμα των βαρών από των εξερχόμενων ακμών από αυτόν τον κόμβο (οι ακμές που φεύγουν από τον κόμβο C_j).

$$\text{Κεντρικότητα κόμβου } (C_i) = IN(C_i) + OUT(C_i) \quad (2),$$

Όπου $IN(C_i)$ είναι το άθροισμα των απόλυτων τιμών των βαρών που εισέρχονται στον κόμβο C_i , $i \neq j$, και το $OUT(C_i)$ είναι το άθροισμα των απόλυτων τιμών των βαρών που εξέρχονται από τον κόμβο C_i , προς τους υπόλοιπους κόμβους C_j , $i \neq j$.

$$IN(C_i) = \sum_{j=1}^n |e_{ji}| \quad (3),$$

$$OUT(C_i) = \sum_{j=1}^n |e_{ij}| \quad (4),$$

Αθροίζοντας τα βάρη μπορούμε να βρούμε την απόλυτη τιμή για κάθε βάρος και έτσι να υπολογίσουμε την θετική ή αρνητική αιτιότητα που δίνει και τη σημαντικότητα σε κάθε κόμβο. Κόμβοι που έχουν μεγάλο βαθμό κεντρικότητας, που είναι αρκετά σημαντικοί, χρειάζονται προσοχή σε οποιαδήποτε ανάλυση του ΑΓΧ που θα προκύψει. [1, 20]

3. Έμμεση και Ολική συνέπεια μεταξύ δύο κόμβων: Ο Kosko [1] περιγράφει μια μέθοδο για εύρεση της έμμεσης και ολικής συνέπειας μεταξύ δύο κόμβων σε ένα μοντέλο ΑΓΧ. Έστω αν υπάρχουν δύο κόμβοι C_i και C_j και οι οποίοι συνδέονται με m διαφορετικά μονοπάτια σχέσεων αιτιότητας. Ένα μονοπάτι μεταξύ του κόμβου C_i και C_j γράφεται ως εξής: $(i, k_{l1}, k_{l2}, \dots, k_{ln}, j)$ για $1 \leq l \leq m$.

Το $I_l(C_i, C_j)$ συμβολίζει την έμμεσες συνέπειες του κόμβου C_i στον κόμβο C_j μέσω του μονοπατιού σχέσεων αιτιότητας, l , και υπολογίζεται ως εξής:

$$I_l(C_i, C_j) = \min\{e(C_p, C_{p+1}) : (p, p+1) \in (i, k_{l1}, k_{l2}, \dots, k_{ln}, j)\} \quad (5),$$

όπου p και $p+1$ συνεχόμενα μονοπάτια.

Το $T(C_i, C_j)$ συμβολίζει την ολικές συνέπειες του κόμβου C_i στον κόμβο C_j μέσω των m μονοπατιών που τους ενώνουν, και υπολογίζεται ως εξής:

$$T(C_i, C_j) = \max(I_l(C_i, C_j)) \quad (6),$$

Άρα η έμμεση συνέπεια από τον κόμβο C_i στον κόμβο C_j δίνεται στην ουσία από το μικρότερο σε τιμή βάρος της συγκεκριμένης σχέσης αιτιότητας στο μονοπάτι μεταξύ των κόμβων C_i και C_j και η ολική συνέπεια είναι η μεγαλύτερη τιμή έμμεσης συνέπειας μετά από τον υπολογισμό όλων των έμμεσων συνεπειών για όλα τα πιθανά μονοπάτια που υπάρχουν από τον κόμβο C_i στον C_j . [1, 20]

4. Βαθμός Ισορροπίας ενός ΑΓΧ: Ο βαθμός ισορροπίας καθορίζει κατά πόσο ένα μοντέλο ΑΓΧ είναι ισορροπημένο ή όχι. Ένας ΑΓΧ δεν είναι ισορροπημένος όταν βρεθούν δύο μονοπάτια μεταξύ δύο ίδιων κόμβων που δημιουργούν σχέσεις αιτιότητας με δύο διαφορετικά πρόσημα, αλλιώς ο ΑΓΧ είναι ισορροπημένος. Για τον υπολογισμό της ισορροπίας γίνεται χρήση του αριθμού των θετικών και αρνητικών ημίκυκλων. Ημίκυκλος σε ένα γράφημα θεωρείται ένας οποιοσδήποτε κύκλος που δημιουργείται χωρίς να λαμβάνουμε υπόψη τις κατευθύνσεις βελών των ακμών. Θετικός ημίκυκλος είναι ο κύκλος στον οποίο αν υπάρχουν βέλη με αρνητικό πρόσημο βαρών είναι ζυγού αριθμού. Αρνητικός ημίκυκλος είναι ο κύκλος στον οποίο υπάρχει μονός αριθμός βελών με αρνητικό πρόσημο βαρών. Ακόμα λαμβάνουμε υπόψη και το μέγεθος του κάθε ημίκυκλου. Ο βαθμός ισορροπίας του ΑΓΧ υπολογίζεται ως εξής: [20]

$$\hat{\alpha} = \frac{\sum m P_m f(m)}{\sum m T_m f(m)} \quad (7), \quad \hat{\alpha} = \frac{\sum m P_m f(m)}{\sum m N_m f(m)} \quad (8),$$

Όπου,

P_m : ο αριθμός των θετικών ημικύκλων μεγέθους m

N_m : ο αριθμός των αρνητικών ημικύκλων μεγέθους m

T_m : ο αριθμός όλων των ημικύκλων μεγέθους m

$f(m)$: μία αυξητική συνάρτηση όπως:

$$f(m) = \frac{1}{m} \text{ ή } f(m) = \frac{1}{m^2} \text{ ή } f(m) = \frac{1}{2^m}$$

2.5 Δυναμική Ανάλυση ΑΓΧ

Μια σημαντική πτυχή του μοντέλου του ΑΓΧ είναι η δυνατότητα στη χρήση του ως εργαλείο λήψης αποφάσεων αλλά και ως εργαλείο για πρόβλεψη μελλοντικών καταστάσεων. Με τη δυναμική ανάλυση του ΑΓΧ περιγράφονται οι αλλαγές που επιφέρουν οι αλληλεπιδράσεις των κόμβων σε κάθε βήμα εκτέλεσης (στη πάροδο του χρόνου). Δεδομένης μιας αρχικής κατάστασης, η οποία αντιπροσωπεύεται με το διάνυσμα τιμών για τους κόμβους- έννοιες του συστήματος, αλλά και του δυσδιάστατου πίνακα των βαρών, το μοντέλο του ΑΓΧ μπορεί να προσομοιώσει την

εξέλιξη της αρχικής κατάστασης μέσα στο χρόνο και να προβλέψει τη μελλοντική συμπεριφορά του συστήματος. Το δίκτυο του ΑΓΧ θα τρέχει σύμφωνα με τον αλγόριθμο μέχρι να καταλήξει σε ένα διάνυσμα με τις τελικές σταθερές τιμές καταστάσεων των κόμβων (Ατελικό). Υπάρχει περίπτωση να καταλήξει σε πολλά διανύσματα τιμών καταστάσεων, όταν παρουσιαστεί το φαινόμενο του limit cycle ή chaotic attractor. Από το τελικό (ή τελικά) διάνυσμα με τις τιμές των καταστάσεων των κόμβων μπορούμε να εξάγουμε πληροφορίες για τις αλλαγές που επέφερε στο σύστημα η συγκεκριμένη αρχική κατάσταση με την οποία έγινε η προσομοίωση. [10, 20]

Αρχή

Βήμα 1: $t = 0$

Διάβασε το αρχικό διάνυσμα τιμών καταστάσεων κόμβων A_0 , $t=0$.

Βήμα 2: Δώσε τον πίνακα με τα βάρη των διασυνδέσεων μεταξύ των κόμβων W .

Βήμα 3: $t = t + 1$

Βήμα 4: Υπολόγισε την καινούργια τιμή καταστάσεως κάθε κόμβου στον ΑΓΧ $A_i(t)$, με μία από τις εξισώσεις που αναφέρονται πιο κάτω στο υποκεφάλαιο 2.5.1

Βήμα 5: Εφαρμόζουμε τη συνάρτηση μεταφοράς στο διάνυσμα με τις νέες τιμές καταστάσεων των κόμβων $f(A(t))$, (βλέπε πιο κάτω, υποκεφάλαιο 2.5.2).

Βήμα 6: Ελέγχουμε εάν ($A(t) == A(t-1)$) ή ($t == \text{μέγιστος αριθμός επαναλήψεων}$) τότε ο αλγόριθμος ΣΤΑΜΑΤΑ, αλλιώς

Πήγαινε στο Βήμα 3.

Πίνακας 2.5.1 Αλγόριθμος προσομοίωσης των αλληλεπιδράσεων του ΑΓΧ

Η μεταβλητή t στον πιο πάνω αλγόριθμο είναι στην ουσία ένας μετρητής των επαναλήψεων που χρειάζεται ο ΑΓΧ για να συγκλίνει σε μια συμπεριφορά, είτε δηλαδή να συγκλίνει σε κατάσταση ισορροπίας, είτε σε ένα κύκλο επαναλήψεων είτε σε χαοτική συμπεριφορά. Ο συμβολισμός $A_i(t)$ είναι η τιμή καταστάσεως του κόμβου i στη χρονική επανάληψη t , ενώ το $A_i(t-1)$ είναι η τιμή καταστάσεως του κόμβου i (C_i) στη χρονική επανάληψη $t-1$, δηλαδή η τιμή της κατάστασης του κόμβου i (C_i) στην

προηγούμενη χρονική επανάληψη. Το $A(t)$ αναφέρεται στο διάνυσμα των τιμών των καταστάσεων όλων των κόμβων. [10,20]

Σε κάθε επαναληπτικό βήμα του αλγορίθμου όπως βλέπουμε οι κόμβοι αλληλεπιδρούν μεταξύ τους επηρεάζοντας ο ένας τον άλλο στο βαθμό που τους επιτρέπουν τα βάρη που τους ενώνουν. Η νέα τιμή του κόμβου, σύμφωνα με την εξίσωση ανανέωσης που παρουσιάζεται στον αλγόριθμο, προκύπτει με τον πολλαπλασιασμό της τιμής των κόμβων που τον επηρεάζουν και του αντίστοιχου βάρους που ενώνει τους κόμβους αυτούς. [10]

Μία νέα προσέγγιση του αλγορίθμου για την αλλαγή των τιμών των οντοτήτων προτάθηκε στο άρθρο [2]. Ο αλγόριθμος που προτείνεται διαφέρει λίγο από τον καθιερωμένο φορμαλισμό των ΑΓΧ που περιγράφηκε πιο πάνω. Βασίζεται στην ιδέα ότι η ανατροπή της κατάστασης ενός συστήματος οφείλεται στις αλλαγές που συμβαίνουν στα επίπεδα ενεργοποίησης των κόμβων. Τα δύο σημεία στα οποία διαφέρει αυτή η προσέγγιση είναι ότι τα βάρη καθορίζονται με διαφορετικό τρόπο και δεν γίνεται χρήση συνάρτησης μεταφοράς, όπως γίνεται στον αλγόριθμο που περιγράψαμε.

Έστω ότι έχουμε ένα σύστημα ΑΣΧ με n αλληλοσχετιζόμενους κόμβους (παραμέτρους), οι οποίοι είναι συνδεδεμένοι με σχέσεις αιτιότητας, τις οποίες θα τις καλούμε ευαισθησία. Κάθε ευαισθησία συσχετίζει τις αλλαγές της κατάστασης ενεργοποίησης του κόμβου C_j ως προς τις αλλαγές της κατάστασης ενεργοποίησης του κόμβου C_i και ορίζεται ως ακολούθως:

$$s_{ij} \equiv \frac{\partial C_j}{\partial C_i} \quad (9),$$

Συνεπώς, η ολική συγκεντρωτική αλλαγή στην κατάσταση ενεργοποίησης του κόμβου C_j λόγω των αλλαγών στους κόμβους C_i (όλοι οι υπόλοιποι κόμβοι που επηρεάζουν τον κόμβο C_j) υπολογίζεται ως ακολούθως:

$$\delta C_j = \sum_{i, i \neq j}^n \frac{\partial C_j}{\partial C_i} \delta C_i = \sum_{i, i \neq j}^n S_{ij} \cdot \delta C_i \quad (10),$$

όπου,

C_i, C_j είναι η δύναμη(τιμή) ενεργοποίησης του κόμβου που μας ενδιαφέρει.

S_{ij} , είναι η ευαισθησία (βάρος). Είναι το μέτρο που μας λέει πόσο πολύ μια αλλαγή στην τρέχουσα τιμή κατάστασης του κόμβου C_i επηρεάζει τις αλλαγές του κόμβου C_j .

Σε διακριτό χρόνο η πιο πάνω εξίσωση γίνεται ως εξής:

$$C_j(t + \delta t) = C_j(t) + \delta C_j = C_j(t) + \sum_{i=1}^n S_{ij} (\delta C_i) \quad (11),$$

Και σε επαναληπτική μορφή υλοποιείται ως ακολούθως:

$$C_j(k + 1) = C_j(k) + \sum_{i, i \neq j}^n S_{ij} (C_i(k) - C_i(k - 1)) \quad (12),$$

Όπου k , είναι ο μετρητής επαναλήψεων.

Ο αλγόριθμος τερματίζει όταν παρουσιαστεί μια από τις τρεις πιο κάτω καταστάσεις:

- Σύγκλιση σε κατάσταση ισορροπίας (steady state): το μοντέλο του ΑΓΧ σταθεροποιείται σε ένα συγκεκριμένο διάνυσμα τιμών καταστάσεων κόμβων, δηλαδή όταν οι τιμές καταστάσεων των κόμβων σταματούν να αλλάζουν σε κάθε βήμα επανάληψης και μένουν σταθερές.
- Σύγκλιση σε κύκλο επαναλήψεων (limit cycle): δηλαδή όταν μια ακολουθία από διανύσματα τιμών καταστάσεων των κόμβων επαναλαμβάνεται σε συγκεκριμένο αριθμό επαναλήψεων η ίδια.
- Χαστική συμπεριφορά (chaotic attractor): όταν οι τιμές των καταστάσεων των κόμβων παρουσιάζουν αστάθεια, δηλαδή αλλάζουν συνεχώς με μη καθοριστικό τρόπο. [10, 20]

Η ικανότητα για πρόβλεψη του μοντέλου του ΑΓΧ είναι χρήσιμη για να απαντηθούν ερωτήματα Εάν – τι? και γίνεται μέσω της δυναμικής αυτής ανάλυσης. Δηλαδή τι θα ακολουθήσει αν πραγματοποιηθεί ένα συμβάν. Θα απαντηθεί μετά το τέλος της ορθή αλυσίδωση, η οποία περιλαμβάνει μια σειρά συνεχόμενων πολλαπλασιασμών των διανυσμάτων τιμών καταστάσεων με τον πίνακα βαρών και μετασχηματισμό του

αποτελέσματος. Οι αρχικές τιμές καταστάσεων τιμών για τους κόμβους αλλά και ο πίνακας βαρών που συνδέουν τους κόμβους καθορίζεται από τους ειδικούς όπου στη συνέχεια δίνονται ως είσοδοι στον αλγόριθμο προσομοίωσης των αλληλεπιδράσεων των κόμβων και αφού τρέξει, θα καταλήξει σε μια συμπεριφορά που αποτελεί την κατάσταση του συστήματος στο μέλλον. [20]

Κάποιοι κόμβοι στο μοντέλο του ΑΓΧ αποκαλούνται κόμβοι πολιτικής (policy nodes). Αυτοί οι κόμβοι κλειδώνονται σε μια συγκεκριμένη τιμή καταστάσεως, έτσι όταν αφήσουμε το δίκτυο του ΑΓΧ να τρέξει σύμφωνα με τον αλγόριθμο, οι τιμές των καταστάσεων των κόμβων που έχουμε κλειδώσει δεν αλλάζουν και παραμένουν σταθερές. Το κλείδωμα αυτό των τιμών καταστάσεως των κόμβων γίνεται εάν ο κόμβος αναπαριστά μια έννοια η οποία γνωρίζουμε εκ των προτέρων ότι δεν πρόκειται να αλλάξει κατά τη περίοδο της προσομοίωσης. [20]

2.5.1 Συναρτήσεις Ανανέωσης Τιμών των Κόμβων

Στον αλγόριθμο προσομοίωσης αλληλεπιδράσεων των κόμβων σε κάθε βήμα της επανάληψης υπολογίζεται νέα τιμή για κάθε κόμβο A_i στο μοντέλο του ΑΓΧ. Υπάρχουν διάφορες υπολογιστικές συναρτήσεις που χρησιμοποιούνται στις διάφορες υλοποιήσεις για την ανανέωση των τιμών αυτών. Όλες όμως βασίζονται στο γεγονός ότι η νέα τιμή A_i για τον κάθε κόμβο C_i εξαρτάτε από τις τιμές των κόμβων που ξεκινούν ακμές και δείχνουν προς αυτόν αλλά και από το βάρος της ακμής που τους ενώνει. Πιο κάτω είναι μερικές από τις συναρτήσεις ανανέωσης τιμών κόμβων:

$$A_i(t) = f\left(\sum_{j=1, j \neq i}^n (A_j(t-1) \cdot W_{ji})\right) \quad (13),$$

$$A_i(t) = f\left(\sum_{j=1, j \neq i}^n (A_j(t-1) \cdot W_{ji}) + A_i(t-1)\right) \quad (14),$$

$$A_i(t) = f\left(\sum_{j=1, j \neq i}^n (k_1 A_j(t-1) \cdot W_{ji} + k_2 A_i(t-1))\right) \quad (15),$$

$$Ai(t)_{(l,k)} = f\left(\sum_{j(l,k)=1, j(l,k) \neq i}^{n(l,k)} (Aj(t-1)_{(l,k)} Wij(t-1)_{(l,k)})\right) - T_{d(l,k)} Ai(t-1)_{(l,k)}$$

(16),

Όπου

$Ai(t)$: η τιμή του κόμβου Ci στην επανάληψη t

$Ai(t-1)$: η τιμή του κόμβου Ci στην επανάληψη $t-1$

$Aj(t-1)$: η τιμή του κόμβου Cj στην επανάληψη $t-1$, και ο οποίος συνδέεται με τον κόμβο Ci

n : ο αριθμός όλων των κόμβων Cj που έχουν ακμή που ξεκινά από αυτούς και δείχνει στον Ci

Wji : το βάρος της ακμής που συνδέει τον κόμβο Cj με τον κόμβο Ci (η ακμή ξεκινά από τον κόμβο Cj και καταλήγει στον κόμβο Ci)

$k1$: ένας συντελεστής που ρυθμίζει κατά πόσο θα επηρεάσει τη νέα τιμή του κόμβου Ci , $Ai(t)$, οι τιμές των κόμβων Cj που συνδέονται μαζί του και έχουν ακμή με βάρος και κατεύθυνση προς τον Ci .

$k2$: ένας συντελεστής που ρυθμίζει κατά πόσο θα επηρεάζει τη νέα τιμή του κόμβου $Ai(t)$ η προηγούμενη τιμή του ιδίου κόμβου (δηλαδή η τιμή που είχε στην προηγούμενη επανάληψη)

$f()$: είναι μια συνάρτηση μεταφοράς. Συναρτήσεις μεταφοράς που υπάρχουν θα αναφερθούν στο υποκεφάλαιο 2.5.2, πιο κάτω.

$Ai(t)_{(l,k)}$: είναι η τιμή ενεργοποίησης του κόμβου Ci στην επανάληψη t στο επίπεδο l στον ΑΓΧ k .

$Aj(t)_{(l,k)}$: είναι η τιμή ενεργοποίησης του κόμβου Cj στην επανάληψη t στο ίδιο επίπεδο l στον ίδιο ΑΓΧ k .

$Td_{(l,k)}$: είναι ο συντελεστής εξασθένησης (decay factor) ο οποίος αφαιρεί ένα ποσοστό από την τιμή ενεργοποίησης του κόμβου στην προηγούμενη επανάληψη έτσι ώστε να εξασθενήσει η επιρροή του πάνω στην παρούσα τιμή ενεργοποίησης του.

2.5.2 Συναρτήσεις Μεταφοράς

Οι συναρτήσεις μεταφοράς χρησιμοποιούνται στον αλγόριθμο προσομοίωσης των αλληλεπιδράσεων για να μετασχηματίζουν το αποτέλεσμα του αθροίσματος του πολλαπλασιασμού των τιμών ενεργοποίησης των κόμβων που επηρεάζουν τον κόμβο του οποίου υπολογίζεται η νέα του τιμή με το αντίστοιχο βάρος της ακμής που τους συνδέει. Αυτό γίνεται για να μεταφέρεται η πραγματική τιμή κάθε μεταβλητής του μοντελοποιημένου συστήματος στο διάστημα μεταξύ $[0,1]$ ή $[-1,1]$. Η επιλογή της κατάλληλης συνάρτησης γίνεται ανάλογα σε ποιο διάστημα θέλουμε να περιορίσουμε τα βάρη. Αν θέλουμε να περιορίζονται στο διάστημα $[0,1]$ τότε μπορεί να χρησιμοποιηθεί η σιγμοειδής συνάρτηση ενώ αν θέλουμε να περιορίζονται μεταξύ $[-1,1]$ μπορούμε να χρησιμοποιήσουμε την $\tanh$. [20, 25]

1. Sigmoid: $f(x) = \frac{1}{1 + e^{-\lambda x}}$ (17),

2. $f(x) = \tanh(x)$ (18),

3. Bivalent: $f(x) = \begin{cases} 0, & x \leq 0, \\ 1, & x \geq 0 \end{cases}$ (19),

4. Trivalent: $f(x) = \begin{cases} -1, & x \leq -0.5, \\ 0, & -0.5 < x < 0.5, \\ 1, & x \geq 0.5, \end{cases}$ (20),

2.6 Διάφορες χρήσεις και εφαρμογές του μοντέλου

Ενώ αρχικά τα μοντέλα των ΑΓΧ προτάθηκαν για εφαρμογές στον τομέα των κοινωνικών επιστημών, όπως για παράδειγμα στην εξέταση πολιτικό-οικονομικών ζητημάτων της Κύπρου [2], στη συνέχεια χρησιμοποιήθηκαν σε ευρύτερο πεδίο εφαρμογών. Παραδείγματα εφαρμογών στα οποία έγινε χρήση μοντέλων ΑΓΧ περιλαμβάνουν ανάλυση ηλεκτρικών κυκλωμάτων, διαχείριση σφάλματος σε καταναεμημένα δίκτυα, μοντελοποίηση και ανάλυση των παραγόντων για την απόδοση

των επιχειρήσεων, στην μάθηση με τη βοήθεια ηλεκτρονικού υπολογιστή όπου καθιστούν ικανό τον η/υ να ελέγχει εάν οι μαθητές έχουν κατανοήσει το μάθημα τους ή όχι, στην μοντελοποίηση πολιτικών υποθέσεων της Νοτίου Αφρικής, στη μοντελοποίηση εικονικών κόσμων. [25] Ακόμα χρησιμοποιούνται στην ανάλυση των παραγόντων υποστήριξης της στρατηγικής CPRF που χρησιμοποιείται από τις επιχειρήσεις σε αλυσίδες τροφοδοσίας [15], στην μοντελοποίηση και ανάλυση παραγόντων για την υποστήριξη της διαχείρισης της δασικής περιοχής Situ Gede που βρίσκεται στην περιοχή Bogor της δυτική Ινδονησία [16], στην μοντελοποίηση των παραγόντων που επηρεάζουν την πορεία της ανάπτυξης έργων λογισμικού [23]. Επίσης αποτελούν βοηθητικό εργαλείο στην λήψη αποφάσεων αφού γίνεται χρήση τους στην μοντελοποίηση συστημάτων διαδικασιών ελέγχου [3]. Ενώ μπορεί ακόμα να χρησιμοποιηθεί και ως βοηθητικό εργαλείο στη διάγνωση ή πρόβλεψη ασθενειών στην ιατρική αφού υπάρχει αναφορά για χρήση του στη διαφορική διάγνωσης της ειδικής γλωσσικής διαταραχής (SLI – specific language impairment) με τις γλωσσικές διαταραχές του αυτισμού και της δυσλεξίας. [22] Αυτά είναι μερικά παραδείγματα της χρήσης τους ενώ στη βιβλιογραφία αναφέρονται πολλά άλλα.

2.7 Αδυναμία μοντέλου ΑΓΧ

Τα περισσότερα μοντέλα ΑΓΧ εάν όχι όλα έχουν αναπτυχθεί από ειδικούς βασισμένα στη γνώση και την εμπειρία τους γύρω από τον τομέα του συστήματος που μοντελοποιείται. Αυτή η διαδικασία ανάπτυξης ενός μοντέλου ΑΓΧ έχει πολλά μειονεκτήματα και αδυναμίες. Καταρχήν απαιτείται η ύπαρξη ειδικού που να έχει γνώσεις γύρω από τον τομέα του συστήματος που μοντελοποιείται αλλά την ίδια στιγμή πρέπει να γνωρίζει και τον φορμαλισμό των ΑΓΧ. Η διαδικασία της ανάπτυξης ενός μοντέλου ΑΓΧ, μπορεί να χρειαστεί να επαναληφθεί πολλές φορές αλλά και να γίνουν αρκετές προσομοιώσεις του μοντέλου μέχρι να αναπτυχθεί ένα κατάλληλο μοντέλο για το σύστημα. Ακόμα σε μεσαίου μεγέθους μοντέλα είναι δύσκολο να εξασφαλιστεί ικανοποιητική απόδοση λόγω του μεγάλου αριθμού παραμέτρων που εμπλέκονται στην ανάπτυξη τους (π.χ. βάρη). Επίσης στις περιπτώσεις όπου τα μοντέλα ΑΓΧ αναπτύσσονται από ομάδα ειδικών λαμβάνονται και άλλοι παράμετροι υπόψη, όπως ο συντελεστής αξιοπιστίας για τον κάθε ειδικό και αυτό προσθέτει επιπλέον πολυπλοκότητα στη διαδικασία. Ένα δεύτερο μειονέκτημα που έχει η

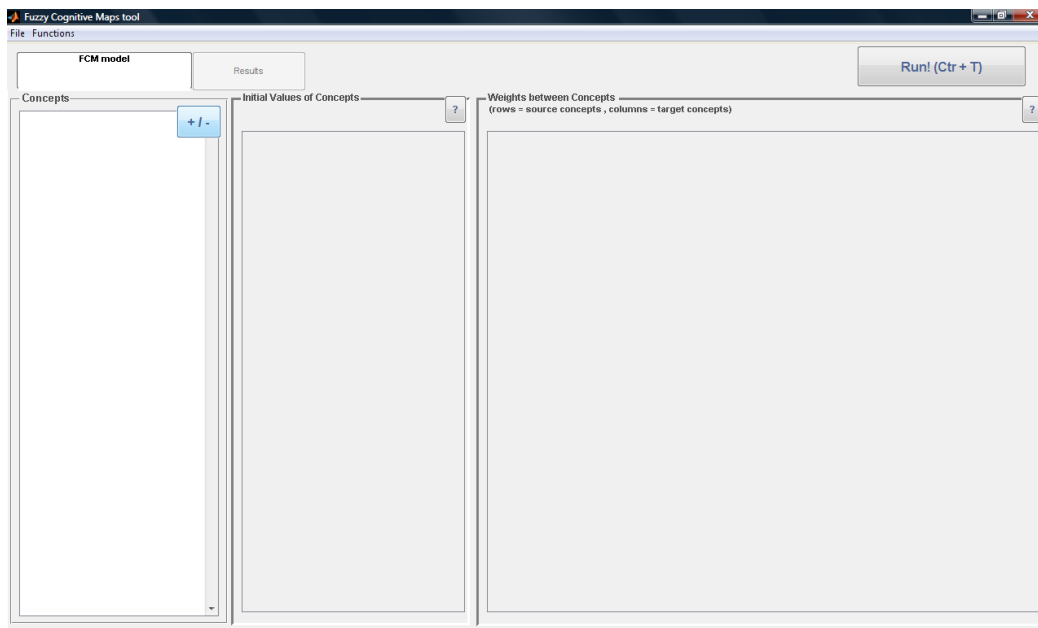
συγκεκριμένη διαδικασία ανάπτυξης των μοντέλων είναι ότι βασίζονται στην ανθρώπινη γνώση. Αυτό υπονοεί ότι η ανάπτυξη του μοντέλου είναι υποκειμενική και λόγω του ότι βασίζονται στην ανθρώπινη κατανόηση είναι δύσκολο να αποτιμηθεί η ορθότητα και η ακρίβεια του μοντέλου με δίκαιο τρόπο. Δύσκολα μπορούν να δικαιολογηθούν ερωτήματα όπως: «Πιστεύω ότι η σχέση αυτή είναι πιο ισχυρή από 0.5, γιατί επέλεξες αυτήν την τιμή για να την εκφράσεις?». Ακόμα με αυτή την προσέγγιση ανάπτυξης είναι πολύ πιθανόν οι ειδικοί να μην καθορίσουν τους κατάλληλους παράγοντες αλλά και να παραλείψουν σχέσεις ανάμεσα σε αυτούς. Σε μεγάλα και σύνθετα προβλήματα που απαιτούν καθορισμό μεγάλο αριθμό κόμβων και συνδέσεων η διαδικασία ανάπτυξης είναι αρκετά δύσκολη. [10, 25]

2.8 Περιγραφή εργαλείου κατασκευής μοντέλου ΑΓΧ και εκτέλεσης δυναμικής ανάλυσης του

Για την κατασκευή διαφόρων μοντέλων ΑΓΧ αλλά και εκτέλεση διαφόρων σεναρίων για δυναμική ανάλυση τους έγινε χρήση ενός εργαλείου το οποίο είναι υλοποιημένο στην γλώσσα προγραμματισμού MATLAB. Το εργαλείο αυτό μέσω ενός φιλικού, εύκολου και εύχρηστου περιβάλλοντος διαπροσωπείας παρέχει στον χρήστη διάφορες επιλογές τις οποίες θα περιγράψουμε πιο κάτω.

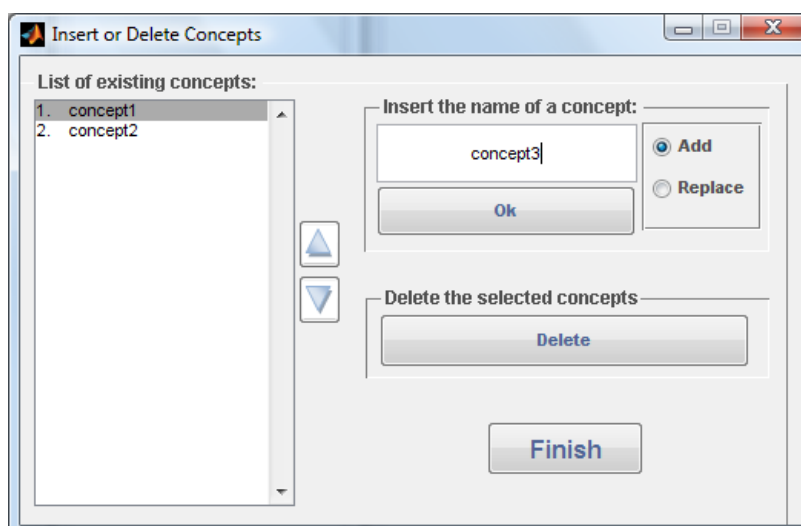
Στο σχήμα 2.7.1 παρουσιάζεται η κύρια οθόνη του εργαλείου, η οποία χωρίζεται σε τρία τμήματα, τους κόμβους, τις αρχικές τιμές των κόμβων και τα βάρη των ακμών ανάμεσα στους κόμβους.

Για την εκτέλεση της δυναμικής ανάλυσης του μοντέλου του ΑΓΧ χρειάζεται να κατασκευαστεί το μοντέλο του ΑΓΧ και άρα θα πρέπει πρώτα να καθοριστούν οι κόμβοι, στη συνέχεια η αρχική κατάσταση που περιγράφει το μοντέλο και τέλος ο πίνακας με τα βάρη των ακμών που συνδέουν τους κόμβους μεταξύ τους. Ο χρήστης αρχικά θα πρέπει να προσθέσει τους κόμβους που περιγράφουν το μοντέλο. Πατώντας το κουμπί +/- που βρίσκεται στην πάνω αριστερή άκρη του πρώτου τμήματος, εμφανίζεται μια δεύτερη οθόνη η οποία φαίνεται στο σχήμα 2.7.2

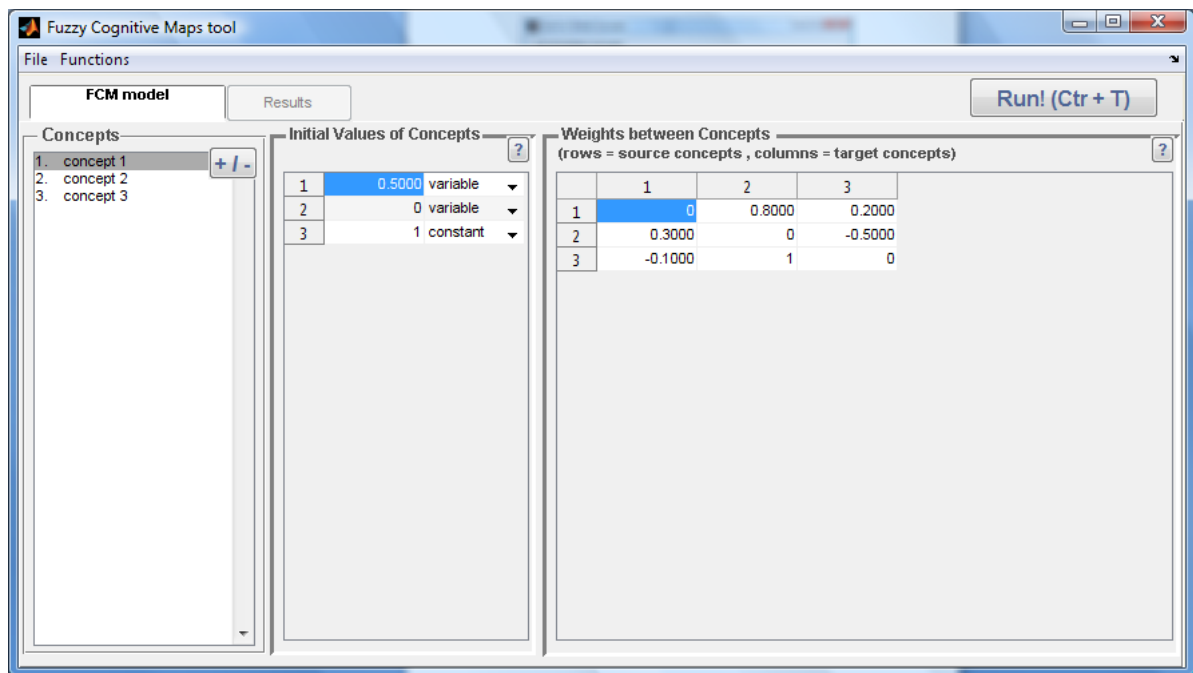


Σχήμα 2.7.1 Κύρια οθόνη εργαλείου εκτέλεσης δυναμικής ανάλυσης ΑΓΧ

Μέσω αυτής της οθόνης (σχήμα 2.7.2) ο χρήστης επιλέγει να προσθέσει, ή να αντικαταστήσει κάποιο κόμβο αφού δώσει το όνομα που περιγράφει την οντότητα που αναπαριστά ο συγκεκριμένος κόμβος. Στα αριστερά αυτής της οθόνης παρουσιάζεται η λίστα με τους κόμβους που έχει προσθέσει στο μοντέλο και ανάλογα μπορεί να αλλάξει την σειρά των κόμβων αλλά και να επιλέξει κάποιον για να τον διαγράψει. Πατώντας το κουμπί Finish ολοκληρώνεται η προσθήκη των κόμβων του μοντέλου και επιστρέφουμε στην αρχική οθόνη η οποία είναι ελαφρώς διαφοροποιημένη και παρουσιάζεται στο σχήμα 2.7.3.



Σχήμα 2.7.2 Οθόνη για πρόσθεση, διαγραφή, αντικατάσταση κόμβων στο μοντέλο ΑΓΧ



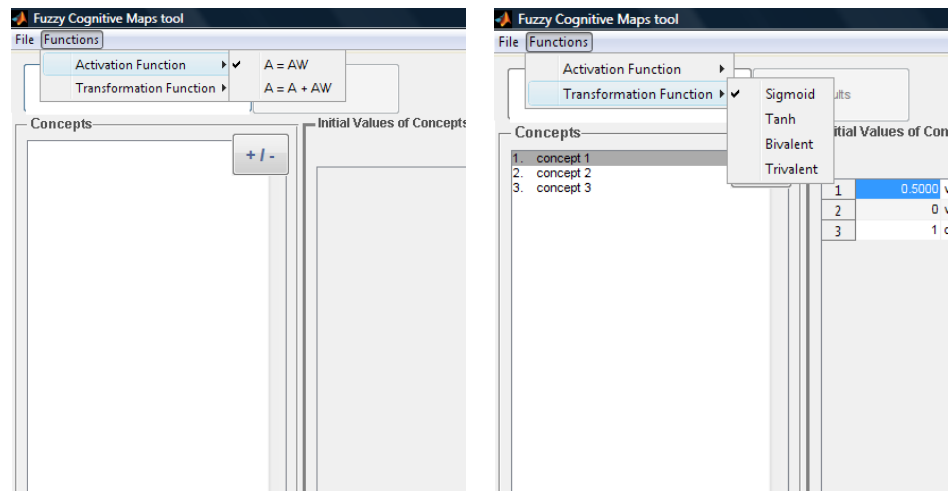
Σχήμα 2.7.3 Κύρια οθόνη που εδώ παρουσιάζει τους κόμβους, τα αρχικά επίπεδα ενεργοποίησης τους, και τον πίνακα βαρών

Αφού ο χρήστης καθορίσει τους κόμβους, μπορεί στη συνέχεια στο δεύτερο τμήμα της πιο πάνω οθόνης να δώσει τις αριθμητικές τιμές για τις τιμές ενεργοποίησης της αρχικής κατάστασης των κόμβων και ανάλογα να τις κλειδώσει ή όχι αν γνωρίζει ότι μιας οντότητας η τιμή της δεν αλλάζει. Οι τιμές αυτές παίρνουν τιμές στο διάστημα $[0,1]$. Στο τρίτο τμήμα ο χρήστης προσδιορίζει τα βάρη των ακμών, για τις διάφορες διασυνδέσεις μεταξύ των κόμβων. Οι γραμμές αντιπροσωπεύουν τους κόμβους από τους οποίους ξεκινά η ακμή(πηγή) και οι στήλες αντιπροσωπεύουν τους κόμβους στους οποίους καταλήγει μια ακμή(στόχος).

Τέλος για την εκτέλεση του αλγορίθμου της δυναμικής ανάλυσης ο χρήστης έχει τη δυνατότητα από το μενού Functions να επιλέξει την συνάρτηση με την οποία θα γίνει η ανανέωση των τιμών ενεργοποίησης των κόμβων. Δίνεται η επιλογή μεταξύ των συναρτήσεων (13) και (14) που ο τύπος τους αναφέρεται στο υποκεφάλαιο 2.5.1.

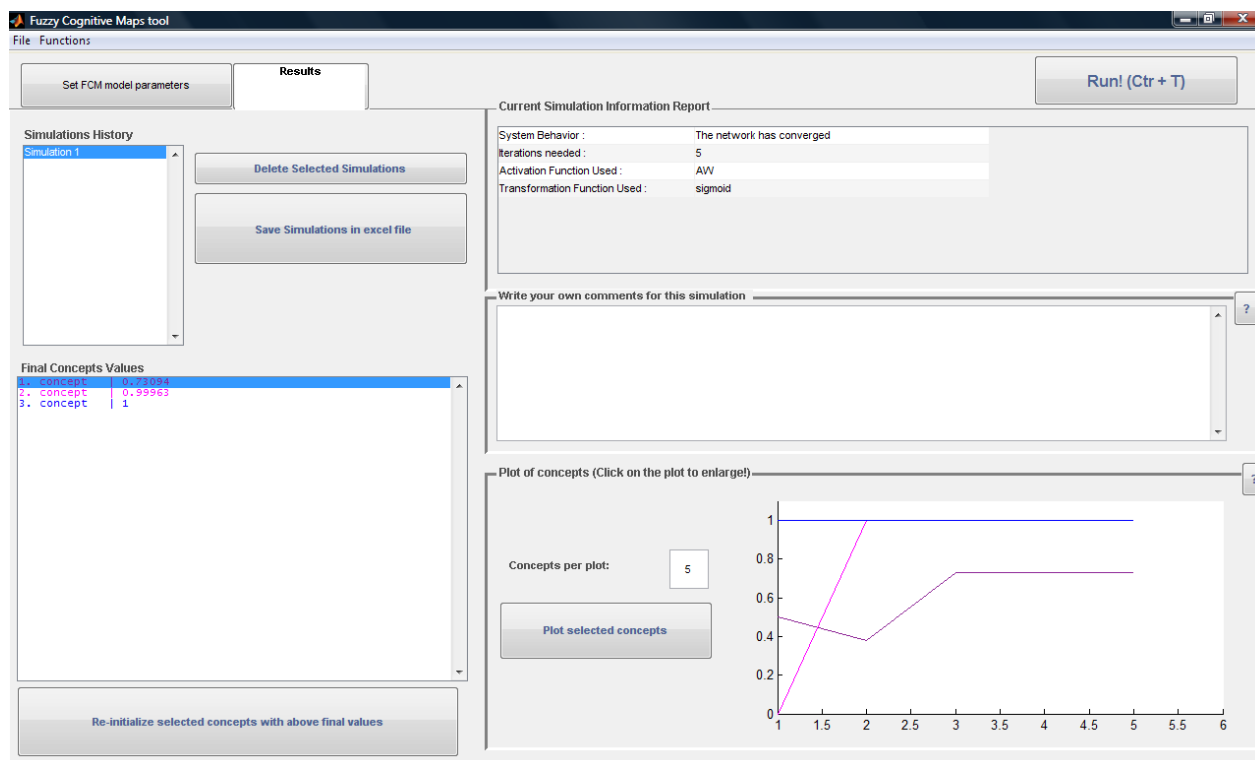
Ακόμα έχει την επιλογή να επιλέξει συνάρτηση μεταφοράς μεταξύ των τεσσάρων συναρτήσεων (17), (18), (19), (20) τους οποίους ο τύπος τους αναφέρθηκε στο υποκεφάλαιο 2.5.2.

Στο σχήμα 2.7.4 παρουσιάζονται οι επιλογές των συναρτήσεων ανανέωσης τιμών ενεργοποίησης των κόμβων και των συναρτήσεων μεταφοράς που δίνει το εργαλείο.



Σχήμα 2.7.4 Οθόνες που παρουσιάζουν την επιλογή συναρτήσεων ανανέωσης τιμών ενεργοποίησης των κόμβων και των συναρτήσεων μεταφοράς

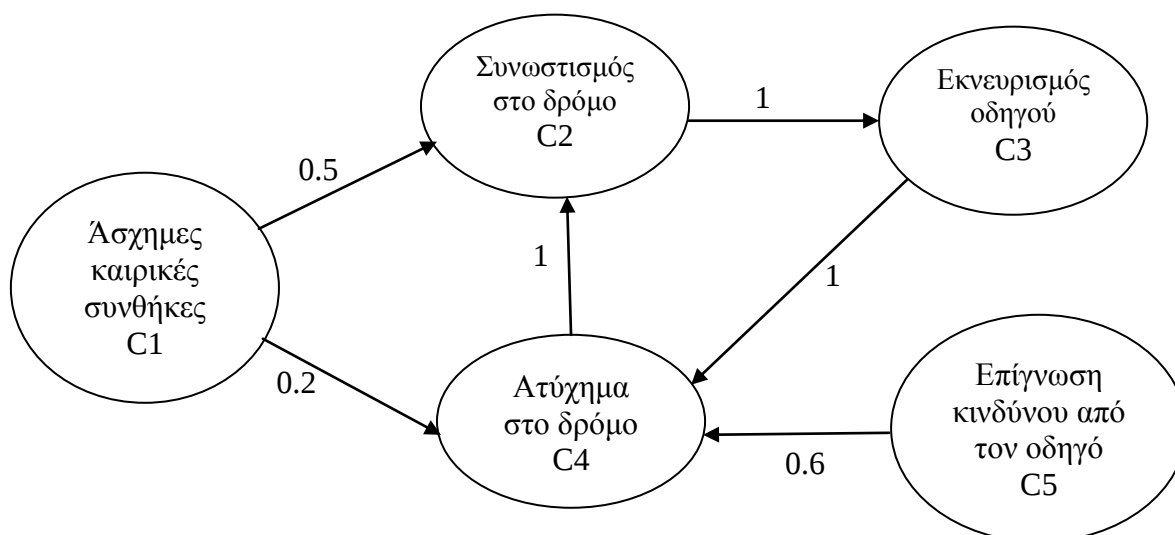
Στο σχήμα 2.7.5 παρουσιάζεται η οθόνη με τα αποτελέσματα της δυναμικής εκτέλεσης του μοντέλου του ΑΓΧ. Για το ίδιο μοντέλο μπορούν να εκτελεστούν πολλές προσομοιώσεις για διαφορετικά σενάρια κάθε φορά. Στο simulation history μένουν τα αποτελέσματα των διάφορων προσομοιώσεων. Το Current Simulation Information Report δίνει μια μικρή αναφορά για την συμπεριφορά του συστήματος για παράδειγμα, αν έχει συγκλίνει σε ένα σταθερό σημείο ισορροπίας, ή έχει χαοτική συμπεριφορά ή αν παρουσιάζεται το φαινόμενο του κύκλου επαναλήψεων. Αν το δίκτυο συγκλίνει σε σταθερό σημείο ισορροπίας δίνεται και ο αριθμός των επαναλήψεων που έτρεξε ο αλγόριθμος μέχρι να συγκλίνει. Οι Final Concept values είναι οι τελικές τιμές των κόμβων μετά το τέλος των επαναλήψεων. Ακόμα το εργαλείο δίνει και σε μορφή γραφικής παράστασης τα αποτελέσματα των προσομοιώσεων. Τα αποτελέσματα αυτά μπορούν να αποθηκευτούν σε αρχείο στην excel (.xls) για περαιτέρω ανάλυση. Επίσης μπορεί κάποιος να τρέξει τον αλγόριθμο με αρχικές τιμές στους κόμβους τις τελικές τιμές μιας προσομοίωσης. Το εργαλείο δίνει τη δυνατότητα αποθήκευσης του μοντέλου του ΑΓΧ σε μορφή .mat έτσι ώστε ο χρήστης να μπορεί να το ξανανοίξει και να εκτελέσει κι άλλες προσομοιώσεις. Ακόμα ένα μοντέλο που κατασκευάστηκε με το εργαλείο μπορεί να εξαχθεί σε αρχείο excel (.xls), αλλά ακόμη και να εισαχθεί ένα μοντέλο από excel (.xls) αρχείο φτάνει να είναι γραμμένο με συγκεκριμένο τρόπο.



Σχήμα 2.7.5 Οθόνη με τα αποτελέσματα της δυναμικής εκτέλεσης του μοντέλου του ΑΓΧ

2.9 Παράδειγμα για επεξήγηση και κατανόηση του μοντέλου του ΑΓΧ

Το πρόβλημα που μοντελοποιεί ο ΑΓΧ είναι τις συνθήκες ενός δρόμου σε ώρα αιχμής.



Σχήμα 2.8.1 Μοντέλο ΑΓΧ για συνθήκες στο δρόμο σε ώρα αιχμής

Οι συνθήκες ενός δρόμου σε ώρα αιχμής ρυθμίζονται από τους 5 πιο κάτω παράγοντες (έννοιες – χαρακτηριστικά):

1. Οι άσχημες καιρικές συνθήκες
2. Ο συνωστισμός στο δρόμο
3. Ο εκνευρισμός των οδηγών
4. Η ύπαρξη ατυχήματος στο δρόμο
5. Η επίγνωση του κινδύνου από τον οδηγό

Αυτοί οι πέντε παράγοντες αποτελούν τους κόμβους του μοντέλου που αναπαριστώνται από τους κύκλους με τις λεκτικές περιγραφές για το κάθε ένα στο εσωτερικό του.

Οι ειδικοί δίνουν ένα αρχικό διάνυσμα με τιμές ενεργοποίησης στους κόμβους του μοντέλου. Το αρχικό αυτό διάνυσμα αντικατοπτρίζει όσο πιο αντικειμενικά γίνεται την πραγματική κατάσταση των συνθηκών ενός δρόμου σε ώρα αιχμής.

Έστω δίνεται το πιο κάτω αρχικό διάνυσμα τιμών ενεργοποίησης:

$$A0 = [0.2, 0.9, 0.7, 0, 1]$$

Το αρχικό διάνυσμα ερμηνεύεται ως εξής:

1. Δεν υπάρχουν άσχημες καιρικές συνθήκες, ο καιρός είναι σχετικά πολύ καλός (C1=0.2)
2. Υπάρχει αυξημένος συνωστισμός στον δρόμο (C2=0.9)
3. Οι οδηγοί είναι αρκετά εκνευρισμένοι (C3=0.7)
4. Δεν υπάρχει κανένα ατύχημα στο δρόμο (C4=0)
5. Οι οδηγοί έχουν πλήρη επίγνωση του κινδύνου που διατρέχουν στο δρόμο. (C5=1)

Οι ειδικοί αφού καθορίσουν τους κόμβους του μοντέλου θα πρέπει να καθορίσουν και τις σχέσεις που τους συνδέουν καθώς και τον βαθμό της κάθε μια σχέσης. Στο παράδειγμα μας οι ειδικοί όρισαν τα εξής βάρη:

1. Βάρος που ενώνει τον κόμβο C1 με τον κόμβο C2 (W12)
2. Βάρος που ενώνει τον κόμβο C2 με τον κόμβο C3 (W23)
3. Βάρος που ενώνει τον κόμβο C3 με τον κόμβο C4 (W34)
4. Βάρος που ενώνει τον κόμβο C5 με τον κόμβο C4 (W54)

5. Βάρος που ενώνει τον κόμβο C1 με τον κόμβο C4 (W14)

6. Βάρος που ενώνει τον κόμβο C4 με τον κόμβο C2 (W42)

Ο αντίστοιχος πίνακας που παρουσιάζει τα βάρη μαζί με τις τιμές τους είναι ο ακόλουθος:

	C1	C2	C3	C4	C5
C1	0.0	0.5	0.0	0.2	0.0
C2	0.0	0.0	1.0	0.0	0.0
C3	0.0	0.0	0.0	1.0	0.0
C4	0.0	1.0	0.0	0.0	0.0
C5	0.0	0.0	0.0	-0.6	0.0

Πίνακας 2.8.2 Πίνακας βαρών του μοντέλου του ΑΓΧ του σχήματος 2.8.1

Οι τιμές των βαρών ερμηνεύονται ως εξής:

1. $W_{12} = 0.5$

Οι άσχημες καιρικές συνθήκες C1 συνεισφέρουν στην δημιουργία συνωστισμού στο δρόμο C2

2. $W_{14} = 0.2$

Οι άσχημες καιρικές συνθήκες (C1) συνεισφέρουν σε **μικρό βαθμό** στο να συμβεί ατύχημα στο δρόμο (C4)

3. $W_{23} = 1$

Όταν δημιουργηθεί συνωστισμός στο δρόμο (C2) τότε **σίγουρα** θα δημιουργηθεί εκνευρισμός στους οδηγούς

4. $W_{34} = 1$

Όταν δημιουργηθεί εκνευρισμός στους οδηγούς (C3) τότε **σίγουρα** θα υπάρξει κάποιο ατύχημα στο δρόμο (C4)

5. $W_{42} = 1$

Όταν συμβεί κάποιο ατύχημα στο δρόμο (C4) τότε **σίγουρα** θα δημιουργηθεί συνωστισμός στο δρόμο (C2)

6. $W_{54} = -0.6$

Όταν οι οδηγοί έχουν επίγνωση του κινδύνου που διατρέχουν στο δρόμο (C5) αυτό συνεισφέρει στο να μην δημιουργηθεί ατύχημα στο δρόμο (C4)

7. Τα υπόλοιπα βάρη έχουν τιμή ίση με 0. $W_{ij}=0$, δηλαδή ότι δεν υπάρχει κάποια σχέση αιτιότητας μεταξύ του κόμβου i και j .

Έχοντας υπόψη τις συναρτήσεις συμμετοχής της μεταβλητής Επιρροής μπορούμε να περιγράψουμε λεκτικά την επιρροή που έχει κάθε κόμβος στους υπόλοιπους άλλους μέσω των τιμών των βαρών του μοντέλου. Βάση του σχήματος 2.3.3, παρουσιάζουμε τις 9 λεκτικές μεταβλητές για κάθε μια από τις συναρτήσεις μεταφοράς για την μεταβλητή Influence(Επιρροή) και το πεδίο αριθμητικής τιμής της.

$\mu_{nvs} \rightarrow$ αρνητική πολύ ισχυρή επιρροή Influence, μεταξύ $[-1,-0.75]$

$\mu_{ns} \rightarrow$ αρνητική ισχυρή επιρροή Influence, μεταξύ $[-1,-0.5]$

$\mu_{nm} \rightarrow$ αρνητική μέση επιρροή Influence, μεταξύ $[-0.75,-0.25]$

$\mu_{nw} \rightarrow$ αρνητική αδύνατη επιρροή Influence, μεταξύ $[-0.5, 0]$

$\mu_z \rightarrow$ μηδενική επιρροή Influence, μεταξύ $[-0.25, 0.25]$

$\mu_{rw} \rightarrow$ θετική αδύνατη επιρροή Influence, μεταξύ $[0,0.5]$

$\mu_{rm} \rightarrow$ θετική μέση επιρροή Influence, μεταξύ $[0.25,0.75]$

$\mu_{rs} \rightarrow$ θετική ισχυρή επιρροή Influence, μεταξύ $[0.5,1]$

$\mu_{rvs} \rightarrow$ θετική πολύ ισχυρή επιρροή Influence $\in [0.75,1]$

1. $W_{12}=0.5$

Περιγράφεται με τη συνάρτηση συμμετοχής μ_{rw} , έχει δηλαδή βαθμό συμμετοχής 1 στο σύνολο της θετικής μέσης επιρροής και μηδέν στα υπόλοιπα σύνολα. Άρα ο κόμβος C1 επηρεάζει μέτρια θετικά τον κόμβο C2.

2. $W_{14} = 0.2$

Περιγράφεται με τις συναρτήσεις συμμετοχής μ_z , σύνολο μηδενικής επιρροής και μ_{rw} , σύνολο θετικής αδύνατης επιρροής, έχοντας μεγαλύτερο βαθμό στο δεύτερο σύνολο. Άρα ο κόμβος C1 επηρεάζει θετικά αλλά σε χαμηλά επίπεδα τον κόμβο C4.

3. $W_{23} = 1$

Περιγράφεται από τη συνάρτηση συμμετοχής μ_{rvs} , με βαθμό συμμετοχής 1 στο σύνολο της θετικής πολύ ισχυρής επιρροής. Άρα μια αλλαγή στον κόμβο

C2 θα επιφέρει τη μέγιστη θετική δυνατή αλλαγή στη κατάσταση του κόμβου C3.

4. W34=1

Περιγράφεται από τη συνάρτηση συμμετοχής μ_{rvs} , με βαθμό συμμετοχής 1 στο σύνολο της θετικής πολύ ισχυρής επιρροής. Άρα μια αλλαγή στον κόμβο C3 θα επιφέρει τη μέγιστη θετική δυνατή αλλαγή στη κατάσταση του κόμβου C4.

5. W42=1

Περιγράφεται από τη συνάρτηση συμμετοχής μ_{rvs} , με βαθμό συμμετοχής 1 στο σύνολο της θετικής πολύ ισχυρής επιρροής. Άρα μια αλλαγή στον κόμβο C4 θα επιφέρει τη μέγιστη θετική δυνατή αλλαγή στη κατάσταση του κόμβου C2.

6. W54 = -0.6

Περιγράφεται από τις συναρτήσεις συμμετοχής μ_{rs} , σύνολο αρνητικά ισχυρής επιρροής και μ_{rm} , σύνολο αρνητικά μέσης επιρροής, με μεγαλύτερο βαθμό συμμετοχής στο σύνολο της αρνητικά ισχυρής επιρροής. Άρα ο κόμβος C5 επηρεάζει αρνητικά σε ψηλά επίπεδα τον κόμβο C4.

Υπολογισμός χαρακτηριστικών στατικής ανάλυσης AGX:

1. Πυκνότητα: Βάση της εξίσωσης (1) του υποκεφαλαίου 2.4 υπολογίζεται ότι είναι: $D = \frac{6}{5^2} = 0.24$. Στο παράδειγμα μας έχουμε 5 κόμβους, άρα $N=5$ και 6 σχέσεις αιτιότητας $R=6$.

2. Υπολογισμός Κεντρικότητας κόμβων:

Με βάση την εξίσωση (2) του υποκεφαλαίου 2.4 και τις εξισώσεις (3) και (4) υπολογίζονται τα εξής:

IN(C1)=0	OUT(C1)= 0.5 + 0.2 =0.7	CEN(C1)=0+0.7=0.7
IN(C2)= 0.5 + 1 =1.5	OUT(C2)=1	CEN(C2)=1.5+1=2.5
IN(C3)=1	OUT(C3)=1	CEN(C3)=1+1=2

$IN(C4)= 1 + 0.2 + -0.6 =1.8$	$OUT(C4)=1$	$CEN(C4)=1.8+1=2.8$
$IN(C5)=0$	$OUT(C5)= -0.6 =0.6$	$CEN(C5)=0+0.6=0.6$

Πίνακας 2.8.3 Υπολογισμός $IN()$, $OUT()$ και Κεντρικότητα ($CEN()$) για το μοντέλο του σχήματος 2.8.1

3. Έμμεση και Ολική Συνέπεια μεταξύ δύο κόμβων

Ας πάρουμε για παράδειγμα ότι θέλουμε να υπολογίσουμε την έμμεση και ολική συνέπεια μεταξύ των δύο κόμβων του παραδείγματος του μοντέλου του ΑΓΧ μας, C1 και C3.

Υπάρχουν δύο μονοπάτια που συνδέουν αυτούς τους κόμβους μεταξύ τους:

Το πρώτο μονοπάτι είναι το εξής: $C1 \rightarrow C2 \rightarrow C3$

Το δεύτερο μονοπάτι είναι το εξής: $C1 \rightarrow C4 \rightarrow C2 \rightarrow C3$

Για το πρώτο μονοπάτι η έμμεση συνέπεια υπολογίζεται ως εξής:

$$I1(C1 \text{ και } C3) = \min \{0.5, 1.0\} = 0.5$$

Η έμμεση συνέπεια είναι η ελάχιστη τιμή του βάρους των ακμών που περιλαμβάνονται στο μονοπάτι μεταξύ των κόμβων C1 και C3. Το 0.5 είναι το βάρος της ακμής μεταξύ των κόμβων C1 και C2, ενώ το 1.0 είναι το βάρος της ακμής μεταξύ των κόμβων C2 και C3.

Για το δεύτερο μονοπάτι η έμμεση συνέπεια υπολογίζεται ως εξής:

$$I2(C1 \text{ και } C3) = \min \{0.2, 1.0, 1.0\} = 0.2$$

Το 0.2 είναι το βάρος της ακμής μεταξύ των κόμβων C1 και C4, το 1.0 είναι το βάρος της ακμής μεταξύ των κόμβων C4 και C2 και το βάρος 1.0 είναι το βάρος της ακμής μεταξύ των κόμβων C2 και C3.

Αφού υπολογίσαμε τις έμμεσες συνέπειες των δύο μονοπατιών, η ολική συνέπεια υπολογίζεται ως η μέγιστη τιμή από τις δύο έμμεσες συνέπειες ως εξής:

$$T(C1 \text{ και } C3) = \max \{ I1(C1 \text{ και } C3), I2(C1 \text{ και } C3) \} = \{0.5, 0.2\} = 0.5$$

4. Βαθμός Ισορροπίας ΑΓΧ

Για να υπολογίσουμε το βαθμό ισορροπίας του ΑΓΧ χρειάζεται να βρούμε τον αριθμό των θετικών και των αρνητικών ημικυκλών που υπάρχουν στο γράφημα μας. Στο παράδειγμα του μοντέλου του ΑΓΧ έχουμε τρεις συνολικά ημικύκλους οι οποίοι είναι οι εξής:

1. $C1 \rightarrow C2 \rightarrow C4 \rightarrow C1$
2. $C2 \rightarrow C3 \rightarrow C4 \rightarrow C2$
3. $C1 \rightarrow C2 \rightarrow C3 \rightarrow C4 \rightarrow C1$

Οι δύο πρώτοι ημικύκλοι έχουν μήκος $m=3$, ενώ ο τρίτος έχει μήκος $m=4$. Οι τρεις αυτοί ημικύκλοι είναι θετικοί αφού σε κανέναν από αυτούς δεν υπάρχει ακμή με αρνητικό πρόσημο βάρους.

Έτσι με τη βάση της εξίσωσης (7) που ο τύπος της αναφέρθηκε στο υποκεφάλαιο 2.4 υπολογίζεται:

$$\hat{\alpha} = \frac{2 * \frac{1}{3} + 1 * \frac{1}{4}}{2 * \frac{1}{3} + 1 * \frac{1}{4}} = 1$$

Όπου,

p_m ο αριθμός των θετικών ημικύκλων με μέγεθος $m=3$ και μέγεθος $m=4$,

$p_3=2$, $p_4=1$

t_m είναι ο αριθμός όλων των ημικύκλων στο γράφο με μέγεθος $m=3$ και $m=4$,

$t_3=2$ και $t_4=1$,

Για την συνάρτηση f χρησιμοποιείται η συνάρτηση $f(m)=1/m$

Το δίκτυο έχει βαθμό ισορροπίας ίσο με 1, άρα είναι ισορροπημένο.

Δυναμική εκτέλεση του μοντέλου του ΑΓΧ

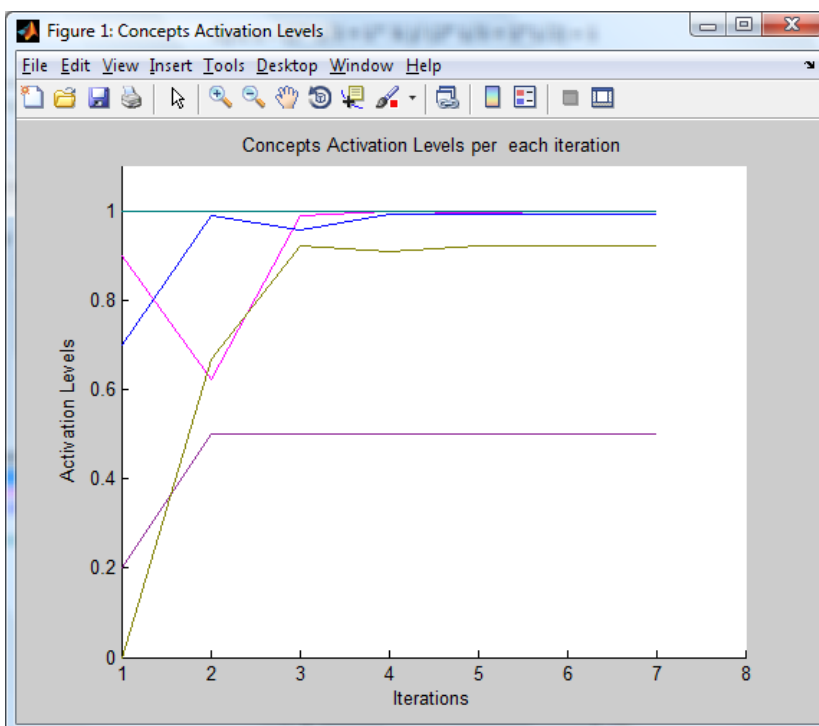
Η δυναμική εκτέλεση του μοντέλου του ΑΓΧ για το συγκεκριμένο πρόβλημα που περιγράφουμε μας βοηθά να απαντήσουμε σε ερωτήματα όπως, ‘Τι θα συμβεί αν ο καιρός είναι κακός?’, ή ‘Τι θα συμβεί αν οι οδηγοί δεν έχουν επίγνωση του κινδύνου στο δρόμο?’, κλπ.

Δίνοντας ως αρχική κατάσταση το διάνυσμα που περιγράψαμε πιο πάνω θέλουμε να δούμε τι θα γίνει στην περίπτωση που ο καιρός είναι σχετικά καλός, υπάρχει αυξημένος συνωστισμός στο δρόμο, οι οδηγοί είναι αρκετά εκνευρισμένοι, αλλά έχουν πλήρη επίγνωση του κινδύνου στο δρόμο και δεν υπάρχει κανένα ατύχημα στο δρόμο. Σημειώνουμε ότι η τιμή της μεταβλητής για την επίγνωση του κινδύνου από τους οδηγούς κλειδώνεται μιας και ξέρουμε ότι δεν θα αλλάξει.

Αρχικό διάνυσμα τιμών των κόμβων:

$$A0 = [0.2, 0.9, 0.7, 0, 1]$$

Με τη βοήθεια του εργαλείου στη Matlab που έχουμε περιγράψει στο υποκεφάλαιο 2.7 πιο πάνω εκτελούμε τον αλγόριθμο. Το δίκτυο αφού αλληλεπιδράσει τελικά συγκλίνει μετά από 7 κύκλους επανάληψης και τα αποτελέσματα φαίνονται στο σχήμα 2.8.1



Σχήμα 2.8.1 Αποτελέσματα των επιπέδων ενεργοποίησης των κόμβων μετά την αλληλεπίδραση των κόμβων

Τελικό διάνυσμα τιμών των κόμβων:

[0.5, 0.99715, 0.99321, 0.92172, 1]

Από το σχήμα 2.8.1 παρατηρείται ότι μετά την αλληλεπίδραση των κόμβων συγκλίνει σε σταθερό επίπεδο ισορροπίας. Τα τελικά αποτελέσματα δείχνουν ότι ενώ ο καιρός ήταν σχετικά καλός, χειροτερεύει, ο συνωστισμός στον δρόμο εξακολουθεί να υπάρχει το ίδιο και ο εκνευρισμός των οδηγών. Παρατηρούμε όμως ότι η τιμή της μεταβλητής για τα ατυχήματα στο δρόμο, αρχικά ήταν 0, τώρα είναι 0.92. Δηλαδή ότι μπορεί να προκληθούν πάρα πολλά ατυχήματα στο δρόμο όταν οι οδηγοί είναι εκνευρισμένοι και υπάρχει μεγάλος συνωστισμός στο δρόμο παρόλο που οι οδηγοί έχουν πλήρη επίγνωση του κινδύνου.

Final Concepts Values	
1. bad weather	0.5
2. free way congestion	0.99715
3. driver frustration	0.99321
4. road accident	0.92172
5. driver risk awareness	1

Σχήμα 2.8.2 Τελικές τιμές των κόμβων

Κεφάλαιο 3

Αλγόριθμοι Εκμάθησης ΑΓΧ

3.1 Γενικά – Αλγόριθμοι Εκμάθησης ΑΓΧ	37
3.2 Αλγόριθμος Εκμάθησης NonLinear Hebbian Learning για ΑΓΧ	41
3.2.1 Περιγραφή αλγορίθμου NHL	41
3.2.2 Καθορισμός Παραμέτρων αλγορίθμου NHL	44
3.2.3 Συνθήκες Τερματισμού	45
3.2.4 Εφαρμογή αλγορίθμου NHL	47
3.3 Γενετικοί Αλγόριθμοι για εκμάθηση ΑΓΧ	49
3.3.1 Βασική Περιγραφή ΓΑ	49
3.3.2 Μέθοδος εκμάθησης μοντέλου ΑΓΧ με Γενετικούς Αλγορίθμους	50
3.3.3 Στοιχεία περιβάλλοντος ΓΑ –Τελεστές και άλλες παράμετροι	52

Στο κεφάλαιο αυτό αναφέρονται γενικά αλγόριθμοι που έχουν χρησιμοποιηθεί για εκπαίδευση του μοντέλου. Γίνεται αναλυτικότερη περιγραφή του αλγορίθμου Nonlinear Hebbian καθώς και προσέγγισης εκμάθησης του μοντέλου των ΑΓΧ με τη χρήση γενετικών αλγορίθμων.

3.1 Γενικά – Αλγόριθμοι Εκμάθησης ΑΓΧ

Στο υποκεφάλαιο 2.6 αναφέραμε την αδυναμία που έχει το μοντέλο ως προς τον τρόπο που γίνεται η ανάπτυξη του. Τα προβλήματα και οι αδυναμίες αυτές των μοντέλων των ΑΓΧ οδήγησαν στην ανάπτυξη υπολογιστικών μεθόδων για την εκμάθηση του πίνακα των συνδέσεων. Οι διαδικασίες εκπαίδευσης μετασχηματίζουν τον πίνακα βαρών των ΑΓΧ και αυξάνουν την απόδοση και την ευρωστία (robustness) τους. Σκοπός των κανόνων εκπαίδευσης είναι η επιλογή κατάλληλων βαρών για τις διασυνδέσεις μεταξύ των οντοτήτων. Με αυτό τον τρόπο τα μοντέλα των ΑΓΧ αποκτούν χρήσιμα

χαρακτηριστικά όπως η ικανότητα τους να μαθαίνουν μόνο τους μη γραμμικούς χάρτες αλλά και την ικανότητα να γενικεύουν καταστάσεις.

Οι μέθοδοι εκμάθησης των μοντέλων που έχουν προταθεί μπορούν να κατηγοριοποιηθούν στις εξής δύο κατηγορίες [18]:

- Hebbian Learning (Εκμάθηση Hebbian)
- Evolutionary Strategies (Εξελικτικές Στατηγικές) και Genetic Algorithm (Γενετικοί Αλγόριθμοι)

Οι αλγόριθμοι εκμάθησης που έχουν προταθεί και ανήκουν στην πρώτη κατηγορία, είναι συνήθως βασισμένοι σε ιδέες που προέρχονται από το χώρο της εκπαίδευσης των τεχνητών νευρωνικών δικτύων αλλά δεν υπάρχει ικανοποιητική απόδειξη των δυνατοτήτων τους, πλήρη περιγραφή των αλγορίθμων και εφαρμογή τους σε κάποιο πραγματικό πρόβλημα. [10]

Ο αλγόριθμος Hebbian είναι ένα από τους πιο συνηθισμένους αλγόριθμους μη επιβλεπόμενης μάθησης στα τεχνητά νευρωνικά δίκτυα. Τα πιο απλά δίκτυα αποτελούνται από ένα σετ διανυσμάτων εισόδου x , ένα σετ διανυσμάτων εξόδου y και ένα πίνακα βαρών W που τα συνδέει, όπου W_{ij} βάρος που συνδέει τον κόμβο X_i με τον κόμβο Y_j . Η μη επιβλεπόμενη μάθηση έχει ως σκοπό να βρει τις τιμές των βαρών W που θα δίνουν τη βέλτιστη λύση στο πρόβλημα που επιλύει το δίκτυο. Κατά τη διάρκεια της εκπαίδευσης το νευρωνικό δίκτυο δέχεται πολλά πρότυπα εισόδου και τα ταξινομεί αυθαίρετα σε κατηγορίες. Ο κανόνας εκμάθησης Hebbian έχει την ακόλουθη μορφή :

$$\Delta W_{ji} = \eta_k \cdot Y_i (X_j - W_{ji} \cdot Y_i) \quad (21),$$

όπου η , θετικός αριθμός, σταθερά εκπαίδευσης, καθορίζει τον ρυθμό εκπαίδευσης. [10, 11]

Ένας κανόνας εκπαίδευσης βαρών απαιτεί τον καθορισμό και υπολογισμό μιας αντικειμενικής συνάρτησης. Στόχος είναι να βρεθεί ένα σετ βαρών που θα ελαχιστοποιεί/μηδενίζει το σφάλμα της αντικειμενικής συνάρτησης είναι να φθάσει σε

ένα ελάχιστο ή μηδενικό σφάλμα που αντιστοιχεί σε ένα σενάριο βαρών του νευρωνικού δικτύου.

Ο Kosko ήταν αυτός που πρότεινε για πρώτη φορά μεθόδους εκπαίδευσης που βασίζονται στον κανόνα μη επιβλεπόμενης μάθησης Hebbian για την προσαρμογή των βαρών στο μοντέλο των ΑΓΧ. Ο Kosko πρότεινε και ανέπτυξε τον αλγόριθμο εκμάθησης χωρίς επίβλεψη που ονομάζεται Διαφορικός Αλγόριθμος εκμάθησης Hebbian (Differential Hebbian Learning, DHL). Ο αλγόριθμος αυτός προτάθηκε για εκπαίδευση μοντέλων ΑΓΧ αλλά χωρίς μαθηματική διατύπωση και εφαρμογή σε πραγματικές περιπτώσεις προβλημάτων. [8] Στο άρθρο [3] χρησιμοποιείται ο συγκεκριμένος αλγόριθμος για εκμάθηση ΑΓΧ που μοντελοποιεί σύστημα διαδικασίας ελέγχου σε χημικές διαδικασίες.

Ένας άλλος αλγόριθμος εκμάθησης που προτάθηκε είναι ο Ισορροπημένος Διαφορικός Αλγόριθμος Εκμάθησης Hebbian (Balanced Differential Hebbian Learning – BDHL). Ο αλγόριθμος αυτός προσπαθεί να εκπαιδεύσει το μοντέλο του ΑΓΧ από τα δεδομένα και είναι βασισμένος στον Διαφορικό αλγόριθμο εκμάθησης (DHL) που πρότεινε ο Kosko. Συγκεκριμένα ο BDHL είναι μια παραλλαγή του DHL, αλλά καλύτερος από τον DHL στο να εκπαιδεύσει το μοντέλο του ΑΓΧ από πρότυπα και να μοντελοποιήσει μια δεδομένη περιοχή από ότι η κλασσική προσέγγιση. [8,10]

Μια άλλη προσέγγιση για εκπαίδευση των ΑΓΧ, που έχει προταθεί από τον Aguilar είναι τα Προσαρμοσμένα Τυχαία Ασαφή Γνωστικά Δίκτυα (Adaptive Random Fuzzy Cognitive Maps) και η οποία στηρίζεται στη θεωρία των τυχαίων νευρωνικών Δικτύων (Random Neural Networks) για να περιγράψουν και να μοντελοποιήσουν την συμπεριφορά ενός συστήματος.[8]

Δύο πιο πρόσφατες προσεγγίσεις είναι οι αλγόριθμοι AHL και NHL. Ο AHL [7,8] Active Hebbian Learning, λαμβάνει υπόψη τη γνώση και την εμπειρία των ειδικών αφού τροποποιούνται τα αρχικά βάρη του πίνακα που θέτουν οι ίδιοι. Ενώ καθορίζεται και μια σειρά ενεργοποίησης, όπου σε κάθε κύκλο του αλγορίθμου κάποιο κόμβοι ενεργοποιούνται και επηρεάζουν άλλους και γίνεται η αλλαγή βαρών. Ακόμα καθορίζονται και κάποιοι κόμβοι εξόδου τους οποίους μας ενδιαφέρει η τιμή τους

αφού οι τιμές αυτών αναπαριστούν την τελική κατάσταση του συστήματος. Ο Nonlinear Hebbian Learning αλγόριθμος περιγράφεται εκτενέστερα στο επόμενο υποκεφάλαιο καθώς και η προσέγγιση μου για υλοποίησή του σε επόμενο κεφάλαιο. [8,11,13]

Επιπρόσθετα με τις μεθόδους των τεχνητών νευρωνικών δικτύων που βασίζονται στην εκμάθηση χωρίς επίβλεψη έχουν προταθεί μέθοδοι βασισμένες σε εξελικτικές στρατηγικές (Evolutionary Strategies – ES). Αυτές οι μέθοδοι ανήκουν στην δεύτερη κατηγορία που αναφέραμε πιο πάνω.

Μια μεθοδολογία που προτάθηκε είναι αυτή της Βελτιστοποίησης με σμήνος σωματιδίων (Particle Swarm Optimization- PSO). Η μέθοδος αυτή παρέχει μια μέθοδο αναζήτησης η οποία βελτιστοποιεί ένα πρόβλημα βάσει μιας συνάρτησης καταλληλότητας $\varphi()$ με το να διατηρεί και να εξελίσσει ένα σμήνος (πλήθος) από υποψήφιες λύσεις. Η υποψήφια λύση από το σμήνος και από όλες τις γενιές η οποία δίνει την καλύτερη τιμή από τη συνάρτηση καταλληλότητας είναι αυτή που δίνει και την βέλτιστη λύση στο πρόβλημα. Με αυτή τη μέθοδο εκμάθησης δημιουργούνται πολλοί πίνακες βαρών και εξάγεται αυτός που θα οδηγήσει το σύστημα σε επιθυμητές περιοχές σύγκλισης. [9]

Μια άλλη προσέγγιση για εκμάθηση μοντέλων ΑΓΧ η οποία βασίζεται σε Εξελικτικές Στρατηγικές και χρησιμοποιήθηκε και στην εκπαίδευση νευρωνικών δικτύων, δεν λαμβάνει υπόψη την αρχική δομή του μοντέλου των ΑΓΧ που καθορίζεται από τη γνώση των ειδικών. Αντίθετα χρησιμοποιεί σύνολα δεδομένων με καθορισμένα πρότυπα εισόδου και εξόδου έτσι ώστε να προσδιοριστούν οι σχέσεις αιτιότητας που ικανοποιούν μια συνάρτηση καταλληλότητας. [6, 8]

Στο άρθρο [12] προτείνεται ακόμα ένας συνδυασμός εκμάθησης αρχικά με τον αλγόριθμο Nonlinear Hebbian Learning και στη συνέχεια βελτιστοποίηση της λύσης του NHL με Εξελικτικό Αλγόριθμο.

Δύο άλλες προσεγγίσεις που έχουν προταθεί βασίζονται σε γενετικούς αλγορίθμους.

Η πρώτη προσέγγιση εκτελεί προσανατολισμένη από τον τελικό στόχο ανάλυση (goal – oriented analysis), στόχος της δεν είναι να υπολογίσει τον πίνακα βαρών αλλά να βρει την αρχική κατάσταση του πίνακα βαρών η οποία θα οδηγήσει το προκαθορισμένο μοντέλο του ΑΓΧ σε μία κατάσταση ισορροπίας (fixed point attractor) ή σε έναν κύκλο επαναλήψεων (limit cycle). [21]

Η δεύτερη μέθοδος που προτείνεται για την ανάπτυξη του πίνακα βαρών του μοντέλου του ΑΓΧ βασίζεται σε ιστορικά δεδομένα που αποτελούνται από ακολουθίες διανυσμάτων καταστάσεων. Η μέθοδος αυτή χρησιμοποιεί πραγματικού κώδικα γενετικό αλγόριθμο (real coded genetic algorithm RCGA) και αυτό επιτρέπει την εξάλειψη της ανθρώπινης παρέμβασης. Στο δεύτερο υποκεφάλαιο που ακολουθεί θα δούμε αναλυτικά την συγκεκριμένη προσέγγιση εκμάθησης του πίνακα βαρών με γενετικούς αλγορίθμους. [24, 25]

3.2 Αλγόριθμος Εκμάθησης Nonlinear Hebbian Learning για τα Ασαφής Γνωστικούς Χάρτες

3.2.1 Περιγραφή του αλγορίθμου NHL

Η προτεινόμενη διαδικασία προσαρμογής των βαρών των ΑΓΧ με την χρήση του αλγορίθμου Nonlinear Hebbian Learning (NHL), βασίζεται στον μη γραμμικό κανόνα τύπου Hebbian που χρησιμοποιήθηκε ως κανόνας μάθησης σε τεχνητά νευρωνικά δίκτυα. Ο κανόνας αυτός έχει προσαρμοστεί και τροποποιηθεί για την περίπτωση των ΑΓΧ, και ονομάζεται Nonlinear Hebbian Learning. [8, 10, 11, 12]

Αυτός ο αλγόριθμος εκμάθησης βασίζεται στην προϋπόθεση ότι όλοι οι κόμβοι του μοντέλου του ΑΣΧ ενεργοποιούνται ταυτόχρονα σε κάθε βήμα επανάληψης, αλληλεπιδρούν μεταξύ τους και αλλάζουν τις τιμές τους. Κατά την διάρκεια αυτής της διαδικασίας ενεργοποίησης το αρχικό μη μηδενικό βάρος W_{ji} των συσχετιζόμενων κόμβων ανανεώνεται και το τροποποιημένο βάρος $W_{ji}(k)$ παράγεται για το επαναληπτικό βήμα k . [10]

Για τον υπολογισμό της τιμής $A_i(k+1)$ του κόμβου C_i στο βήμα επανάληψης $k+1$, υπολογίζεται η επιρροή των διασυνδεδεμένων κόμβων C_j με αντίστοιχες τιμές A_j στον

συγκεκριμένο κόμβο C_i λόγω της αλλαγής των βαρών $W_{ji}(k)$ στο βήμα επανάληψης k , σύμφωνα με την πιο κάτω εξίσωση [10]:

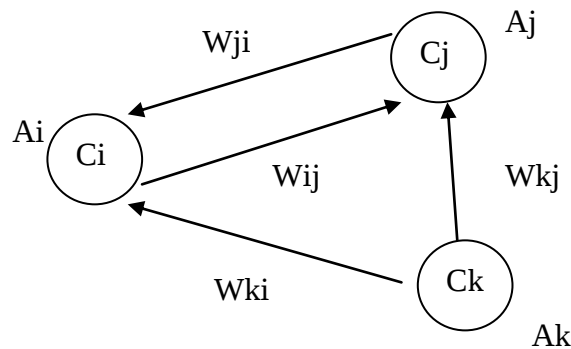
$$A_i(k+1) = f \left(A_i(k) + \sum_{j=1}^N A_j(k) \cdot W_{ji}(k) \right) \quad (22),$$

Έστω για παράδειγμα ότι στο βήμα επανάληψης $k+1$, όπως φαίνεται σχήμα 3.2.1.1 ο κόμβος C_i αλληλεπιδρά στον κόμβο C_j , ο κόμβος C_j αλληλεπιδρά στον κόμβο C_i και ο κόμβος C_k αλληλεπιδρά και στον κόμβο C_i και στον κόμβο C_j . Οι τιμές των κόμβων C_i , C_j και C_k αλλάζουν σύμφωνα με την εξίσωση (1). Αναλυτικά οι νέες τιμές $A_i(k+1)$, $A_j(k+1)$, $A_k(k+1)$ υπολογίζονται ως εξής [10]:

$$A_i(k+1) = f \left(A_i(k) + (A_j(k) \cdot W_{ji}(k) + A_k(k) \cdot W_{ki}(k)) \right)$$

$$A_j(k+1) = f \left(A_j(k) + (A_i(k) \cdot W_{ij}(k) + A_k(k) \cdot W_{kj}(k)) \right)$$

$$A_k(k+1) = f \left(A_k(k) \right)$$



Σχήμα 3.2.1.1 Αλληλεπίδραση τριών κόμβων για υπολογισμό τιμών στο βήμα $k+1$, οι τιμές στους κόμβους και στα βάρη στο σχήμα είναι του βήματος k .

Εκτός από τον υπολογισμό των τιμών των κόμβων σε κάθε επαναληπτικό βήμα γίνεται τροποποίηση των βαρών, δηλαδή των τιμών των διασυνδέσεων μεταξύ των κόμβων. Σε κάθε βήμα επανάληψης πραγματοποιούνται μικρές μεταβολές γύρω από τις τιμές που έχουν τα βάρη στο προηγούμενο επαναληπτικό βήμα, επιδιώκοντας έτσι να οδηγηθεί το σύστημα σε σύγκλιση σε επιθυμητές τιμές. [10]

Αξιοποιώντας τα χαρακτηριστικά του γενικού μη γραμμικού κανόνα εκμάθησης Hebbian για τα νευρωνικά δίκτυα παρουσιάζεται ο μαθηματικό φορμαλισμό που

προτείνεται για το μοντέλο των ΑΓΧ. Ο κανόνας συσχετίζει τις τιμές των κόμβων και τις τιμές των βαρών στο μοντέλο του ΑΓΧ και έχει την πιο κάτω γενική μαθηματική μορφή:

$$\Delta W_{ji} = n \cdot A_i(k) \cdot (A_j(k) - W_{ji}(k) \cdot A_i(k)) \quad (23),$$

όπου ο συντελεστής n είναι ένας πάρα πολύ μικρός θετικός παράγοντας που ονομάζεται παράμετρος ποσοστού εκμάθησης και ο οποίος καθορίζεται χρησιμοποιώντας την πειραματική μέθοδο δοκιμής και λάθους (trial and error) ούτως ώστε να βελτιστοποιήσουμε την τελική λύση. [10]

Η εξίσωση αυτή τροποποιείται και προσαρμόζεται για τους ΑΓΧ και προτείνεται η ακόλουθη μορφή του κανόνα τροποποίησης των βαρών:

$$W_{ji}(k) = \gamma \cdot W_{ji}(k-1) + n \cdot A_j(k-1) \cdot (A_i(k) - A_j \cdot W_{ji}(k-1)) \quad (24),$$

όπου ο συντελεστής n είναι η παράμετρος μάθησης και γ είναι η παράμετρος μείωσης του βάρους. [10]

Η σχέση που προτείνεται δεν προκαλεί μεγάλες τροποποιήσεις γιατί ο όρος $n \cdot A_j(k-1) \cdot (A_i(k) - A_j \cdot W_{ji}(k-1))$ είναι μικρός σε σχέση με τον πρώτο όρο $\gamma \cdot W_{ji}(k-1)$.

Ο κανόνας εκμάθησης των βαρών ερμηνεύεται ως ακολούθως, για τον υπολογισμό της νέας τιμή του βάρους στο βήμα k , λαμβάνεται υπόψη κατά ένα μεγάλο ποσοστό η τιμή του βάρους στο προηγούμενο βήμα του αλγορίθμου η οποία πολλαπλασιάζεται με τον συντελεστή μείωσης του βάρους ($\gamma \cdot W_{ji}(k)$) και κατά ένα μικρότερο ποσοστό τις τιμές των κόμβων στο προηγούμενο βήμα που αλληλεπιδρούν πολλαπλασιασμένες με τον συντελεστή n . [10]

Η τιμή του κάθε κόμβου του ΑΣΧ ανανεώνεται σε κάθε βήμα σύμφωνα με την εξίσωση (22), αφού πρώτα υπολογιστεί η νέα τιμή των βαρών $W_{ji}(k)$ σύμφωνα με την εξίσωση (24).

Στο άρθρο [13] ο αλγόριθμος NHL χρησιμοποιείται και ανανεώνει ταυτόχρονα σε κάθε βήμα επανάληψης όλα τα βάρη του πίνακα βαρών εκτός από τα βάρη της διαγωνίου ενώ στα άρθρα [8] και [11] ανανεώνονται μόνο τα αρχικά μη – μηδενικά βάρη τα

οποία προτάθηκαν από τους εμπειρογνώμονες σύμφωνα με την εξίσωση τροποποίησης των βαρών μέχρι τον τερματισμό του αλγορίθμου. Με τον πρώτο τρόπο όλα τα βάρη αλλάζουν τιμή και αφού αλλάζουν και τα μηδενικά βάρη θέτονται νέες διασυνδέσεις μεταξύ των κόμβων ενώ με τον δεύτερο τρόπο ο αλγόριθμος NHL δεν θέτει νέες διασυνδέσεις μεταξύ των κόμβων και όλα τα μηδενικά βάρη δεν αλλάζουν τιμή και παραμένουν μηδενικά.

Κατά την διάρκεια του σταδίου κατασκευής του ΑΓΧ, οι εμπειρογνώμονες επιλέγουν κάποιους κόμβους που ονομάζονται Κόμβοι Εξόδου Επιθυμητής τιμής (Desired Output Concepts – DOCs). Οι κόμβοι αυτοί αντιπροσωπεύουν κάποια κύρια χαρακτηριστικά του συστήματος που μας ενδιαφέρουν και των οποίων θέλουμε να υπολογίσουμε τις τιμές των καταστάσεων τους, οι οποίες αντικατοπτρίζουν την ολική κατάσταση του συστήματος. Η διάκριση των κόμβων του ΑΓΧ σε κόμβους εισόδου και εξόδου καθορίζεται από την ομάδα των εμπειρογνομώνων για κάθε συγκεκριμένο πρόβλημα. Οι εμπειρογνώμονες επιλέγουν τους κόμβους εξόδους και θεωρούν τους υπόλοιπους ως αρχικούς διεγέρτες ή εσωτερικούς κόμβους του συστήματος. [10]

3.2.2 Καθορισμός παραμέτρων εκμάθησης του αλγορίθμου NHL

Όπως αναφέραμε στην εξίσωση που χρησιμοποιούμε για την εκμάθηση των βαρών υπάρχουν δύο παράμετροι, το n που είναι η παράμετρος ποσοστού εκμάθησης και η παράμετρος γ που είναι ο συντελεστής μείωσης βάρους. Οι τιμές των παραμέτρων αυτών, για κάθε περίπτωση εξασφαλίζουν ότι η διαδικασία εκμάθησης προσαρμόζει τα βάρη ώστε το σύστημα να συγκλίνει σε μια επιθυμητές καταστάσεις ισορροπίας. Για τις τιμές αυτές χρειάζεται να καθοριστεί ένα άνω και κάτω όριο στις τιμές τις οποίες μπορούν να πάρουν αυτές οι παράμετροι. Στα άρθρα [8,11,13] μετά από πειράματα δοκιμής και λάθους προέκυψε ότι τα όρια αυτά καθορίζονται από 0 μέχρι 0.1 για την παράμετρο n , $0 < n < 0.1$ και από 0.9 μέχρι 1 για την παράμετρο γ , $0.9 < \gamma < 1$. Για μεγάλες τιμές του n (μεγαλύτερες του 0.2) η τιμή του βάρους μεταβάλλεται αρκετά και είτε αλλάζει πρόσημο είτε το σύστημα δεν συγκλίνει και συνεχώς ταλαντεύεται. Όσο αφορά τον παράγοντα γ , παρατηρήθηκε ότι όταν αυτός παίρνει μικρές τιμές τα βάρη μεταβάλλονται αρκετά και το σύστημα δεν συγκλίνει, και η προηγούμενη τιμή του βάρους συνεισφέρει ελάχιστα για τον υπολογισμό της νέας τιμής. Όσο μεγαλύτερη

είναι η τιμή του γ , τόσο πιο πολύ λαμβάνεται υπόψη η προηγούμενη τιμή του βάρους για τον υπολογισμό της νέας τιμής. [10]

3.2.3 Συνθήκες τερματισμού του αλγόριθμου

Για τον τερματισμό του αλγορίθμου έχουν προταθεί οι εξής συνθήκες που πρέπει να ικανοποιούνται για να τερματιστεί η διαδικασία.

Στο άρθρο [13] προτείνεται ως κριτήριο τερματισμού η μεγιστοποίηση μιας αντικειμενικής συνάρτησης J , η οποία εξετάζει τις επιθυμητές τιμές των κόμβων εξόδου. Οι τιμές των κόμβων αυτών άλλωστε είναι αυτές που μας ενδιαφέρουν. Η αντικειμενική συνάρτηση J , που προτείνεται για τον NHL ορίζεται ως ακολούθως:

$$J = \sum_{i=1}^l (\text{DOC } i)^2 \quad (25),$$

όπου l είναι ο αριθμός των κόμβων εξόδου.

Μια άλλη συνθήκη τερματισμού που προτείνεται στο άρθρο [8] είναι η ελαχιστοποίηση της τιμής της συνάρτησης F (26) που δίνεται πιο κάτω:

$$F = \sqrt{(\|\text{DOC}_j(k) - T_j\|)^2} \quad (26),$$

Η συνάρτηση αυτή υπολογίζει το τετράγωνο της Ευκλείδειας απόστασης της πραγματικής τιμής DOC_j στο βήμα επανάληψης k και της μέσης τιμής, T_j , των επιθυμητών τιμών για τις τιμές των κόμβων εξόδου (DOC_j) στο επαναληπτικό βήμα k . Ο όρος T_j στην συνάρτηση (26) είναι η μέση τιμή της τιμής που επιθυμούμε να έχουν οι κόμβοι εξόδου (DOC_j).

Για παράδειγμα έστω ότι θέλουμε να υπολογίσουμε την τιμή της συνάρτησης F του κόμβου C_j . Απαιτείται ότι οι κόμβοι εξόδου επιθυμητής τιμής (DOC_j) παίρνουν τιμές στο διάστημα ενός πεδίου $\text{DOC}_j = [T_j \text{ min}, T_j \text{ max}]$. Τότε η μέση τιμή του στόχου DOC_j για τον κόμβο C_j καθορίζεται ως εξής:

$$T_j = \frac{(T_j \text{ min} + T_j \text{ max})}{2} \quad (27),$$

Σε ένα μοντέλο ΑΓΧ, μπορεί να υπάρχουν m DOCs, έτσι για τον υπολογισμό της συνάρτησης F παίρνουμε το άθροισμα των διαφορών των τετραγώνων μεταξύ των τιμών των m DOCs με τις m T 's μέσες τιμές των DOCs και άρα η εξίσωση (27) έχει την ακόλουθη μορφή:

$$F1 = \sqrt{\sum_{j=1}^m (DOC_j(k) - T_j)^2} \quad (28),$$

Ένα ακόμη κριτήριο που προτείνεται στα άρθρα [8,13] για τον τερματισμό του NHL αλγορίθμου βασίζεται στο ότι επιδιώκουμε η διαφορά δύο ακόλουθων τιμών των κόμβων DOC_j για το επαναληπτικό βήμα k να είναι μικρότερη από μια πολύ μικρή τιμή ϵ . Το κριτήριο αυτό έχει την εξής μορφή:

$$F2 = |DOC_j(k+1) - DOC_j(k)| < \epsilon \quad (29),$$

Όπου $DOC_j(k)$ είναι η τιμή του j οστού κόμβου στο επαναληπτικό βήμα k , και το ϵ είναι μια πάρα πολύ μικρή τιμή. Για την ϵ δίνονται τιμές 0.001, 0.002

Ψευδοκώδικας του αλγορίθμου Nonlinear Hebbian Learning
Βήμα 1: Δώσε το αρχικό διάνυσμα της κατάστασης των κόμβων $A0$, τον αρχικό πίνακα βαρών $W0$ και τον πίνακα με τους κόμβους επιθυμητής εξόδου, καθώς και μέγιστο αριθμό επαναλήψεων, $k=0$; Βήμα 2: Καθόρισε τις τιμές των παραμέτρων $n(k)$ και $\gamma(k)$ Βήμα 3: $k=k+1$ (επανάλαβε τα εξής για κάθε βήμα επανάληψης k) Βήμα 4: Ανανέωσε τα μη – μηδενικά βάρη Βήμα 5: Υπολόγισε τις τιμές $A_j(k)$ Βήμα 6: Υπολόγισε τις δύο συνθήκες τερματισμού Βήμα 7: Αν δεν ικανοποιούνται οι συνθήκες τερματισμού, ή $k > \text{μέγιστες επαναλήψεις}$ γύρισε στο βήμα 2 Βήμα 8: Επέστρεψε τον τελικό πίνακα με τα τροποποιημένα βάρη W_{NHL}

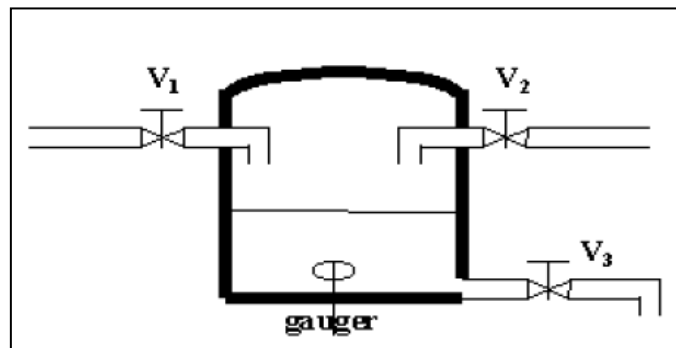
Σχήμα 3.2.3.1: Ψευδοκώδικας του αλγορίθμου Nonlinear Hebbian Learning

3.2.4 Εφαρμογή αλγορίθμου NHL

Ο αλγόριθμος που προτάθηκε στα άρθρα [8,10,11,12] εφαρμόστηκε σε απλά παραδείγματα διαδικασιών ελέγχου. Στο σημείο αυτό θα αναφέρουμε το πρόβλημα διαδικασίας του άρθρου [13] στο οποίο έχουμε μια δεξαμενή με τρεις βαλβίδες και οι οποίες επηρεάζουν την ποσότητα του υγρού στην δεξαμενή. Το σχήμα 3.2.4.1 απεικονίζει το σύστημα. Από τις βαλβίδες 1 και 2 εισέρχονται στην δεξαμενή δύο διαφορετικού τύπου υγρά μέσα στην δεξαμενή και όταν αναμειχτούν αυτά τα υγρά δημιουργείται μια χημική αντίδραση. Ένας αισθητήρας μετρά την συγκεκριμένη βαρύτητα του υγρού που παράγεται μέσα στη δεξαμενή. Όταν η τιμή αυτού του υγρού είναι μεταξύ ενός πεδίου τιμών, (G_{max}) και (G_{min}) σημαίνει ότι έχει παραχθεί το επιθυμητό υγρό στην δεξαμενή. Επίσης υπάρχει ακόμη ένας περιορισμός στο σύστημα για το ύψος του υγρού το οποίο δεν πρέπει να ξεπερνά ένα ανώτατο όριο (H_{max}) ενώ δεν μπορεί να είναι κάτω από ένα κατώτατο όριο (H_{min}). Έτσι ο στόχος του ελέγχου είναι να διατηρούνται οι τιμές των συγκεκριμένων μεταβλητών στα πιο κάτω πεδία τιμών:

$$0.74 \leq G \leq 0.80 \text{ και}$$

$$0.68 \leq H \leq 0.70$$



Σχήμα 3.2.4.1 Απεικόνιση συστήματος απλής διαδικασίας

Τρεις ειδικοί ανέπτυξαν ένα μοντέλο AGX για το πιο πάνω σύστημα. Καθόρισαν τους κόμβους και τις σχέσεις ανάμεσα τους και όρισαν ασαφή βάρη σε αυτές. Το σχήμα του μοντέλου παρουσιάζεται στο σχήμα 3.2.4.2. Το μοντέλο αποτελείται από τους εξής πέντε κόμβους:

Κόμβος 1 – η ποσότητα του υγρού που περιέχεται στη δεξαμενή (εξαρτάτε από την κατάσταση των βαλβίδων 1,2 και 3)

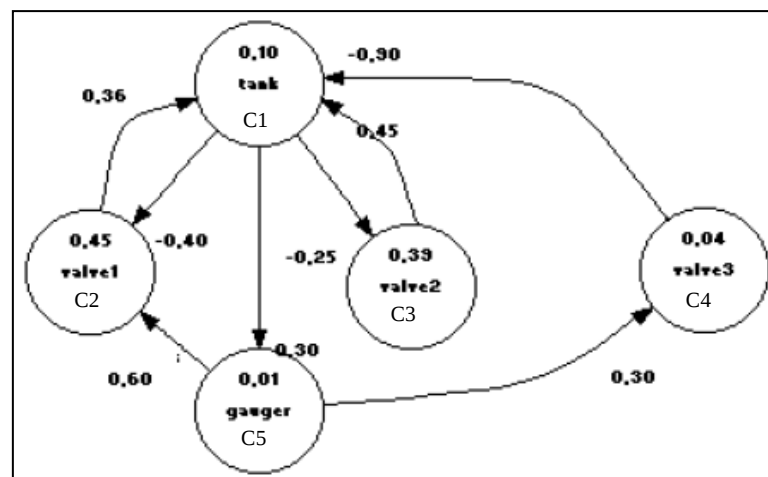
Κόμβος 2 – η κατάσταση της βαλβίδας 1 (κλειστή, ανοικτή, ή μερικώς ανοικτή)

Κόμβος 3 – η κατάσταση της βαλβίδας 2 (κλειστή, ανοικτή, ή μερικώς ανοικτή)

Κόμβος 4 – η κατάσταση της βαλβίδας 3 (κλειστή, ανοικτή, ή μερικώς ανοικτή)

Κόμβος 5 – η βαρύτητα του υγρού στη δεξαμενή.

Για το συγκεκριμένο πρόβλημα οι ειδικοί καθόρισαν τις αρχικές τιμές των κόμβων και των βαρών και οι οποίες παρουσιάζονται στο σχήμα 3.2.4.2. Επίσης ως κόμβοι επιθυμητής τιμής εξόδου καθορίστηκαν οι κόμβοι C1 και C5.



Σχήμα 3.2.4.2 Μοντέλο ΑΓΧ για το σύστημα

Έγινε εκπαίδευση του μοντέλου του ΑΓΧ με τη συγκεκριμένη αρχική κατάσταση και τον πίνακα βαρών που καθορίστηκε από τους έμπειρους και προέκυψε ένας νέος πίνακας βαρών με νέες διασυνδέσεις ανάμεσα στους κόμβους. Στη συνέχεια εκτελέστηκε ο αλγόριθμος αλληλεπίδρασης των κόμβων με τον νέο πίνακα βαρών για 1000 τυχαία διανύσματα αρχικών καταστάσεων των κόμβων. Για όλες τις περιπτώσεις οι τελικές τιμές των κόμβων επιθυμητής τιμής εξόδου C1 και C5 κυμαίνονταν στα πεδία τιμών που αναφέρθηκαν πιο πάνω και τα οποία καθορίστηκαν από τους ειδικούς.

3.3 Γενετικοί Αλγόριθμοι για εκμάθηση ΑΓΧ

3.3.1 Βασική περιγραφή γενετικών αλγορίθμων

Οι γενετικοί αλγόριθμοι αποτελούν μια ξεχωριστή κατηγορία των εξελικτικών αλγορίθμων που είναι ένα υποσύνολο του εξελικτικού υπολογισμού και συγκεκριμένα, ένας αλγόριθμος βελτιστοποίησης βασισμένος στον πληθυσμό του. Οι γενετικοί αλγόριθμοι έχουν εμπνευστεί από τον Δαρβίνο και τη θεωρία του για την βιολογική εξέλιξη, απ' όπου και χρησιμοποιούν ορισμένους μηχανισμούς, όπως την μετάλλαξη, τον ανά-συνδυασμό και την επιλογή. Ο αλγόριθμος ξεκινά με ένα σύνολο από πιθανές λύσεις για το πρόβλημα που έχουμε να επιλύσουμε που καλείται πληθυσμός. Οι πιθανές αυτές λύσεις διαδραματίζουν το ρόλο των χρωμοσωμάτων σε έναν πληθυσμό. Κάποιες λύσεις από έναν πληθυσμό επιλέγονται και χρησιμοποιούνται για τον σχηματισμό ενός νέου πληθυσμού με την ελπίδα ότι ο νέος πληθυσμός θα είναι καλύτερος από τον προηγούμενο. Οι λύσεις που επιλέγονται για να σχηματίσουν νέες λύσεις (απόγονοι) επιλέγονται με βάση την καταλληλότητα τους την οποία καθορίζει μια συνάρτηση καταλληλότητας – οι πιο κατάλληλες είναι πιο πιθανόν να αναπαραχθούν. Η εξέλιξη του πληθυσμού γίνεται με την επανειλημμένη εφαρμογή των πιο πάνω μηχανισμών μέχρις ότου ικανοποιηθεί κάποια συνθήκη (για παράδειγμα ένας αριθμός πληθυσμών ή βελτίωση της καλύτερης λύσης). [30]

Βήματα Βασικού Γενετικού Αλγορίθμου [30]

1. **[Αρχή]** Παραγωγή τυχαίου πληθυσμού με n χρωμοσώματα (κατάλληλες λύσεις για το πρόβλημα)
2. **[Καταλληλότητα]** Αξιολόγηση της καταλληλότητας $f(x)$ του κάθε χρωμοσώματος x μέσα στον πληθυσμό.
3. **[Νέος πληθυσμός]** Δημιουργία νέου πληθυσμού με την επανάληψη των πιο κάτω βημάτων μέχρι να ολοκληρωθεί ο νέος πληθυσμός.
 1. **[Επιλογή]** Επέλεξε δύο χρωμοσώματα γονείς από έναν πληθυσμό ανάλογα με την καταλληλότητα τους (τα καλύτερα σε βαθμό καταλληλότητας, έχουν την μεγαλύτερη πιθανότητα να επιλεγούν)
 2. **[Διασταύρωση]** Με μία πιθανότητα διασταύρωσης, διασταυρώνονται οι γονείς για να σχηματίσουν ένα νέο απόγονο (παιδιά). Εάν δεν γίνει καμία διασταύρωση τότε ο απόγονος είναι ακριβές αντίγραφο των γονιών.

3. **[Μετάλλαξη]** Με μια πιθανότητα μετάλλαξης, μεταλλάσσεται ο νέος απόγονος σε κάθε θέση χρωμοσώματος (locus).
4. **[Αποδοχή]** Τοποθέτησε τον νέο απόγονο μέσα σε ένα νέο πληθυσμό.
4. **[Αντικατάσταση]** Χρησιμοποιούμε τον νέο παραγόμενο πληθυσμό για τις επόμενες εκτελέσεις του αλγόριθμου.
5. **[Έλεγχος]** Εάν η συνθήκη τερματισμού ικανοποιηθεί, ο αλγόριθμος **σταματά**, και επιστρέφει την καλύτερη λύση στην συγκεκριμένη θέση.
6. **[Επανάληψη-Βρόγχος]** Πήγαινε στο βήμα 2

Σχήμα 3.3.1.1 Βήματα Βασικού Γενετικού Αλγορίθμου.

Το Σχήμα 3.3.1.1 που παρουσιάζει τα βήματα του βασικού γενετικού αλγορίθμου είναι πολύ γενικό. Υπάρχουν πολλές παράμετροι που μπορούν να υλοποιηθούν διαφορετικά ανάλογα με το πρόβλημα που έχουμε να επιλύσουμε. Καταρχήν θα πρέπει να κατασκευαστούν τα χρωμοσώματα, άρα απαιτείται ένας τρόπος κωδικοποίησης τους. Ανάλογα με τον τύπο κωδικοποίησης που θα χρησιμοποιηθεί για την κατασκευή των χρωμοσωμάτων εξαρτώνται οι δύο βασικότεροι τελεστές των γενετικών αλγορίθμων που είναι η διασταύρωση και η μετάλλαξη. Ακόμα θα πρέπει να καθοριστεί ο τρόπος με τον οποίο θα γίνεται η επιλογή των χρωμοσωμάτων που στη συνέχεια θα αποτελέσουν τους γονείς για να γίνει η διασταύρωση για την παραγωγή των νέων απογόνων. Η επιλογή μπορεί να γίνει με διάφορους τρόπους αλλά η κύρια ιδέα είναι να επιλέγονται οι καλύτεροι γονείς με την ελπίδα ότι έτσι θα παραχθούν καλύτεροι απόγονοι. Ακόμα για την παραγωγή του νέου πληθυσμού μπορεί τουλάχιστον μια από τις καλύτερες λύσεις από τον παλιό πληθυσμό να επιλεγεί για τον νέο πληθυσμό χωρίς να γίνουν σε αυτήν αλλαγές. [30]

3.3.2 Μέθοδος εκμάθησης μοντέλου ΑΓΧ με Γενετικούς Αλγόριθμους

Στόχος της εκμάθησης του μοντέλου του ΑΓΧ είναι να καθοριστεί ο πίνακας συνδέσεων με τις τιμές των βαρών W . Στην προσέγγιση που προτείνεται στα άρθρα [24,25] χρησιμοποιεί μια επέκταση του Γενετικού Αλγορίθμου που καλείται Real-coded genetic algorithm (RCGA), λόγω της κωδικοποίησης του χρωμοσώματος, για την εκμάθηση του μοντέλου του ΑΓΧ.

Τα δεδομένα για την εκμάθηση είναι ένα σύνολο κόμβων N , μια ακολουθία από τιμές ενεργοποίησης κατάστασης των κόμβων N , $C(t)$, για ένα συγκεκριμένο χρονικό διάστημα $t \in \{0, \dots, t_k\}$ – που αποτελούν τα δεδομένα εισόδου και κάποια συνάρτηση μετασχηματισμού f , η οποία θα ελαχιστοποιεί το λάθος.

Η συγκεκριμένη μέθοδος εκμάθησης με τη βοήθεια γενετικών αλγορίθμων έχει ως σκοπό να εξαλείψει την ανθρώπινη παρέμβαση κατά τη διάρκεια της ανάπτυξης του μοντέλου του ΑΓΧ. Για τη διαδικασία αυτή γίνεται χρήση πληροφοριών που προέρχονται από ιστορικά δεδομένα για τον υπολογισμό του πίνακα συνδέσεων, ο οποίος θα είναι ικανός να μιμείται τα δεδομένα. Τα δεδομένα εισόδου είναι μια ακολουθία καταστάσεων η οποία περιγράφεται από διανύσματα καταστάσεων σε μια συγκεκριμένη χρονική περίοδο. Το σύνολο των διανυσμάτων των καταστάσεων $A(t)$ μια χρονικής περιόδου t αναπαριστά την συμπεριφορά του συστήματος στη πάροδο του χρόνου.

Η κύρια ιδέα ήταν ότι αφού ήταν δεδομένη μια ακολουθία τιμών καταστάσεων ενεργοποίησης των κόμβων για ένα χρονικό διάστημα K , μπορούσαν να ομαδοποιηθούν ανά δύο τα συνεχόμενα διανύσματα καταστάσεων και να προκύψουν $K-1$ διαφορετικά ζευγάρια, που θα είχαν την πιο κάτω μορφή:

$$A(t) \rightarrow A(t+1) \quad \forall t = 0, \dots, K-1 \quad (30),$$

Εάν καθορίσουμε το $A(t)$ ως το αρχικό διάνυσμα τιμών καταστάσεων και το $A(t+1)$ την ανταπόκριση του συστήματος, τότε μπορούν να παραχθούν $k-1$ ζευγάρια της μορφής {αρχικό διάνυσμα, ανταπόκριση συστήματος} από τα δεδομένα εισόδου. Όσο πιο μεγάλος είναι ο αριθμός K τόσο πιο πολλές πληροφορίες για τη συμπεριφορά του συστήματος έχουμε. Με αυτό τον τρόπο δημιουργούνται τα δεδομένα με τα οποία θα εκπαιδευτεί το μοντέλο μας.

Το μοντέλο του ΑΓΧ αφού εκπαιδευτεί θα πρέπει να παράγει την ίδια ακολουθία διανυσμάτων καταστάσεων για το ίδιο διάνυσμα τιμών αρχικής κατάστασης το οποίο καθορίζεται από τα δεδομένα εισόδου. [24, 25]

3.3.3 Στοιχεία περιβάλλοντος Γενετικών Αλγορίθμων –Τελεστές και άλλες παράμετροι

1. Δομή χρωμοσώματος

Το χρωμόσωμα περιέχει πληροφορίες για την λύση που αναπαριστά. Υπάρχουν διάφοροι τρόποι για κωδικοποίηση ενός χρωμοσώματος και αυτό συνήθως εξαρτάτε από το πρόβλημα το οποίο επιλύεται. Ο πιο συνηθισμένος τρόπος κωδικοποίησης είναι μια σειρά από ψηφία 0 και 1. [30]

Στην περίπτωση του προβλήματος της εκμάθησης του πίνακα των συνδέσεων του ΑΓΧ το χρωμόσωμα κωδικοποιείται ως ένα διάνυσμα με αριθμούς κινητής υποδιαστολής (floating-point numbers). [24, 25]

Το μήκος του χρωμοσώματος αντιστοιχεί στον αριθμό των μεταβλητών (βάρη των συνδέσεων) σε ένα δεδομένο πρόβλημα. Κάθε στοιχείο του διανύσματος ονομάζεται γονίδιο. [30]

Σημειώνεται ότι στην προσέγγιση των άρθρων [24,25] οι ακμές επιτρέπονται μόνο μεταξύ διαφορετικών κόμβων, δηλαδή ότι οι κόμβοι δεν έχουν σχέσεις αιτίας αποτελέσματος από και προς τον εαυτό τους. Έτσι ο πίνακας συνδέσεων του μοντέλου του ΑΓΧ μπορεί να περιγραφεί από $N \times (N-1)$ μεταβλητές, όπου N είναι ο αριθμός των κόμβων. Άρα η εκμάθηση του μοντέλου του ΑΓΧ δεν είναι τίποτε άλλο από τον υπολογισμό αυτών των $N \times (N-1)$ μεταβλητών. Κάθε χρωμόσωμα αποτελείται από $N \times (N-1)$ γονίδια τα οποία είναι αριθμοί κινητής υποδιαστολής και ανήκουν στο διάστημα $[-1,1]$, και έχουν την πιο κάτω μορφή:

$$W = [W_{12}, W_{13}, \dots, W_{1N}, W_{21}, W_{23}, \dots, W_{2N}, \dots, W_{NN-1}]^T \quad (31),$$

όπου W_{ij} καθορίζει την τιμή του βάρους της ακμής από τον κόμβο i στον κόμβο j .

Κάθε χρωμόσωμα αποκωδικοποιείται σε ένα υποψήφιο μοντέλο ΑΓΧ αντιγράφοντας τις τιμές των βαρών από το χρωμόσωμα στο αντίστοιχο κελί στον πίνακα συνδέσεων και έτσι προκύπτει ο πίνακας συνδέσεων του μοντέλου του ΑΓΧ.

Ο αριθμός των χρωμοσωμάτων σε ένα πληθυσμό είναι μια σταθερά για κάθε γενιά και καθορίζεται από μία παράμετρο που ονομάζεται μέγεθος_πληθυσμού.

Ένας δεύτερος τρόπος κωδικοποίησης του χρωμοσώματος και αναφέρεται στο άρθρο [19] θα μπορούσε να ήταν στη μορφή δυσδιάστατου πίνακα όπως φαίνεται στο πιο κάτω σχήμα (3.3.3.1):

W11	W12	W13	W14	W15
W21	W22	W23	W24	W25
W31	W32	W33	W34	W35
W41	W42	W43	W44	W45
W51	W52	W53	W54	W55

Σχήμα 3.3.3.1 Κωδικοποίηση χρωματοσώματος σε δυσδιάστατο τετραγωνικό πίνακα

Όπου η θέση $[i,j]$ του πίνακα έχει το βάρος της ακμής που ξεκινά από τον κόμβο C_i και καταλήγει στον C_j .

2. Συνάρτηση Καταλληλότητας

Ένας από τους πιο σημαντικούς παράγοντες που πρέπει να ληφθεί υπόψη για να είναι επιτυχημένη μια εφαρμογή γενετικών αλγορίθμων είναι ο σχεδιασμός μιας συνάρτησης καταλληλότητας η οποία θα πρέπει να είναι κατάλληλη για ένα δεδομένο πρόβλημα. Η συνάρτηση καταλληλότητας είναι ένας τρόπος για να αξιολογούνται τα χρωμοσώματα που αναπαριστούν τις διάφορες υποψήφιες λύσεις και να δίνεται σε αυτές μια τιμή ανάλογα με το πόσο κατάλληλη είναι μια λύση για το δεδομένο πρόβλημα. [30]

Στην περίπτωση της εφαρμογής για εκμάθηση του πίνακα συνδέσεων ενός ΑΓΧ, η συνάρτηση καταλληλότητας περιλαμβάνει μια συνάρτηση σφάλματος που υπολογίζει τη διαφορά ανάμεσα στην ανταπόκριση του συστήματος από το υποψήφιο μοντέλο, με την αντίστοιχη πραγματική ανταπόκριση του συστήματος η οποία είναι γνωστή από τα δεδομένα εισόδου. [24,25]

Για να βρεθεί η ανταπόκριση του συστήματος ενός υποψήφιου μοντέλου ΑΓΧ γίνεται αποκωδικοποίηση του χρωμοσώματος (βάρη ακμών) σε ένα μοντέλο ΑΓΧ και εκτελείται μια επανάληψη στον αλγόριθμο προσομοίωσης αλληλεπιδράσεων με αρχικό διάνυσμα το αρχικό διάνυσμα των δεδομένων εισόδου. Η διαφορά υπολογίζεται για όλα τα $K-1$ {αρχικό διάνυσμα καταστάσεων, ανταπόκριση συστήματος} ζευγάρια.

Η εξίσωση για υπολογισμό της διαφοράς (υπολογισμό του σφάλματος) είναι η ακόλουθη[24,25]:

$$Difference = \frac{1}{(K-1)N} \sum_{t=1}^{K-1} \sum_{n=1}^N |An(t) - \hat{An}(t)|^2 \quad (32),$$

όπου

$A(t) = [A1(t), A2(t), \dots, An(t)]$, η γνωστή (δεδομένα) ανταπόκριση του συστήματος για το $A(t-1)$ αρχικό διάνυσμα τιμών ενεργοποίησης.

$\hat{A}(t) = [\hat{A}1(t), \hat{A}2(t), \dots, \hat{A}n(t)]$, η ανταπόκριση του συστήματος όταν εκτελεστεί ο αλγόριθμος αλληλεπιδράσεων για ένα υποψήφιο μοντέλο ΑΓΧ για το $A(t-1)$ αρχικό διάνυσμα τιμών ενεργοποίησης.

Ο όρος $\frac{1}{(K-1)N}$, είναι παράμετρος για κανονικοποίηση της τιμής της διαφοράς (σφάλματος).

Η εξίσωση υπολογισμού του λάθους χρησιμοποιείται ως βασικό τμήμα για την συνάρτηση καταλληλότητας η οποία έχει την εξής μορφή:

$$Fitness\ function = h(Difference)$$

Όπου h είναι μια βοηθητική συνάρτηση.

Η βοηθητική συνάρτηση h χρησιμοποιείται για δύο κύριους λόγους:

1. Για να διασφαλίσει ότι τα καλύτερα χρωμοσώματα ανταποκρίνονται σε μεγάλες τιμές της συνάρτησης καταλληλότητας.
2. Προσθέτει μη γραμμικότητα έτσι ώστε να ανταμείψει τα χρωμοσώματα που είναι πιο κοντά στην επιθυμητή λύση. [24,25]

Προτείνεται η πιο κάτω βοηθητική συνάρτηση h:

$$h(x) = \frac{1}{(ax+1)} \quad (33),$$

όπου a είναι μια παράμετρος που παίρνει τιμή εμπειρικά, $a = 10000$.

Αρα η συνάρτηση καταλληλότητας καθορίζεται ως εξής:

$$h(Difference) = \frac{1}{(10000 \text{ Difference} + 1)} \quad (34),$$

Η τιμή της συνάρτηση καταλληλότητας κανονικοποιείται σε τιμές που ανήκουν στο πεδίο $[0,1]$, όπου:

- Τα χειρότερα χρωμοσώματα έχουν τιμή καταλληλότητας κοντά στο 0
- Τα ιδανικότερα χρωμοσώματα (κοντά στη λύση) των οποίων τα αποτελέσματα είναι ακριβώς η ίδια ακολουθία διανυσμάτων τιμών ενεργοποίησης με αυτήν των δεδομένων εισόδου έχουν τιμή ίση με 1.

3. Κριτήρια τερματισμού

Τα κριτήρια τερματισμού για την διαδικασία εκμάθησης του μοντέλου του ΑΓΧ λαμβάνουν υπόψη τέσσερα πιθανά σενάρια:

- Η εκμάθηση είναι επιτυχημένη, για παράδειγμα με την προσομοίωση του υποψήφιου μοντέλου ΑΓΧ παίρνουμε μια ακολουθία διανυσμάτων ενεργοποίησης τιμών η οποία είναι ίδια ή ικανοποιητικά κοντινή με τα δεδομένα εισόδου. Το πόσο ικανοποιητικά κοντά είναι με την πραγματική το καθορίζει η τιμή της συνάρτηση καταλληλότητας για το καλύτερο χρωμόσωμα σε κάθε γενιά. Έτσι η διαδικασία εκμάθησης θα πρέπει να τερματίζει όταν η τιμή της συνάρτησης καταλληλότητας φτάσει σε μια τιμή που καλούμε `max_fitness`.
- Μέγιστος χρόνος εκτέλεσης αλγορίθμου: Η εκμάθηση δεν είναι επιτυχημένη, αλλά έχει παραχθεί ένας μέγιστος αριθμός γενεών, ο οποίος καθορίζεται από τη παράμετρο `max_generation`.

- Η καλύτερη λύση δε βελτιώνεται: Εάν η καλύτερη λύση δεν βελτιώνεται άλλο μετά από μια χρονική περίοδο τότε ο αλγόριθμος τερματίζει. (ο χρήστης καθορίζει αυτή τη χρονική περίοδο) [24,25]

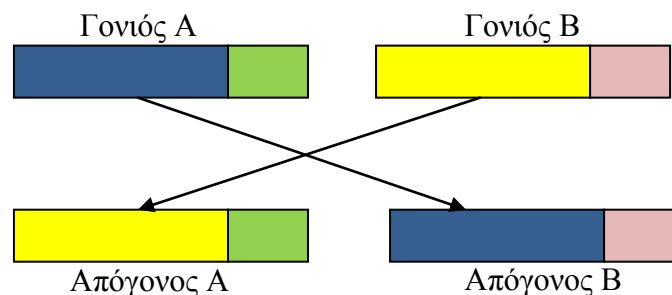
Οι δύο παράμετροι `max_fitness` και `max_generation` καθορίζονται εμπειρικά και θα αναφέρουμε τις τιμές τους στη συνέχεια.

4. Μέθοδος ανά-συνδυασμού γονιδίων (Crossover Operator)

Ο τελεστής διασταύρωσης είναι μια μέθοδος για να ανταλλάσσονται πληροφορίες μεταξύ των χρωμοσωμάτων. Συνδυάζει τα χαρακτηριστικά δυο χρωμοσωμάτων που αποτελούν τους γονείς για τον σχηματισμό δύο νέων χρωμοσωμάτων που αποτελούν τους απογόνους. Αυτό συμβαίνει με την ελπίδα ότι δύο καλά χρωμοσώματα έχουν πιθανότητα να σχηματίσουν δύο καλύτερα χρωμοσώματα.[30]

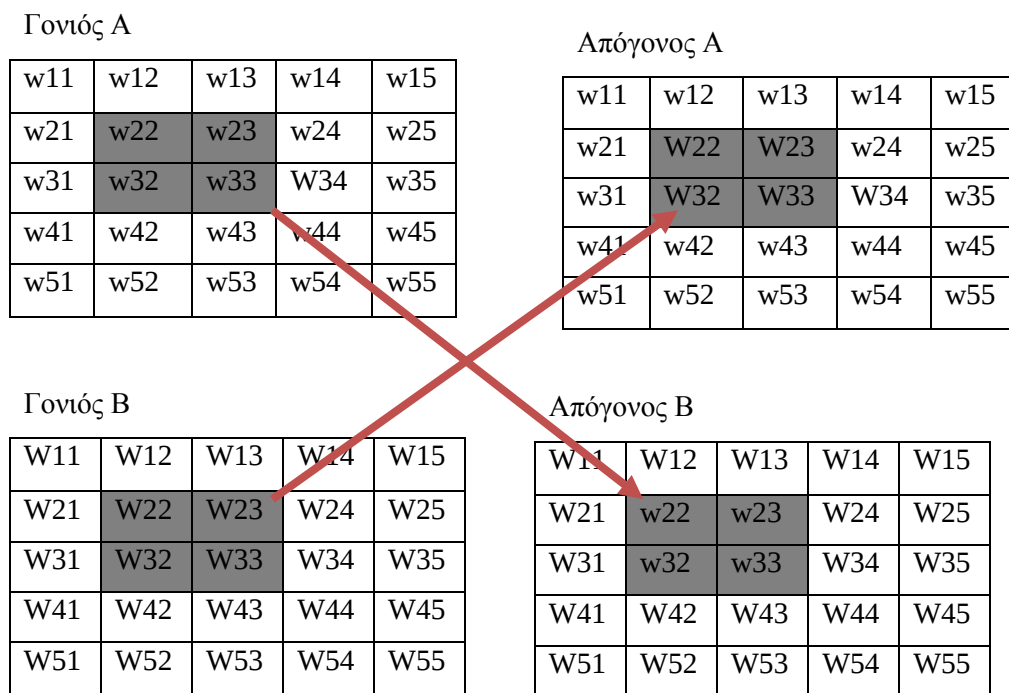
Στην μέθοδο των άρθρων [24,25] χρησιμοποιείται η διασταύρωση ενός σημείου (one – point crossover). Με αυτόν τον τελεστή διασταύρωσης επιλέγεται τυχαία ένα σημείο διασταύρωσης και οτιδήποτε πριν από αυτό το σημείο αντιγράφεται στο χρωμόσωμα απόγονο από τον πρώτο γονιό ενώ οτιδήποτε μετά από αυτό το σημείο αντιγράφεται στο χρωμόσωμα απόγονο από τον δεύτερο γονιό. Με τον ίδιο τρόπο πραγματοποιείται η διασταύρωση δύο σημείων (two –point crossover) αλλά επιλέγονται δύο τυχαία σημεία.

Στο σχήμα 3.3.3.2 φαίνεται η διασταύρωση ενός σημείου από δύο χρωμοσώματα γονείς και ο σχηματισμός δύο νέων απογόνων:



Σχήμα 3.3.3.2 Ανά-συνδυασμός ενός σημείου (crossover) από δύο χρωμοσώματα γονείς και σχηματισμός δύο νέων απογόνων.

Στην περίπτωση που η κωδικοποίηση του χρωμοσώματος δεν είναι αριθμός κινητής υποδιαστολής, αλλά είναι δυσδιάστατος πίνακας ο πιο πάνω τελεστής διασταύρωσης δεν μπορεί να εφαρμοστεί. Ο τρόπος με τον οποίο εκτελείται η διασταύρωση είναι ο εξής: Επιλέγονται δύο αριθμοί στα όρια της στήλης του πίνακα και δύο αριθμοί στα όρια της γραμμής του πίνακα με τυχαίο τρόπο και ανταλλάσσεται μεταξύ των δύο πινάκων η περιοχή που επιλέγεται. Η εφαρμογή του τελεστή όταν η κωδικοποίηση του χρωμοσώματος είναι σε μορφή πίνακα φαίνεται στο σχήμα 3.3.3.3 [19]:



Σχήμα 3.3.3.3 Εφαρμογή τελεστή crossover όταν η κωδικοποίηση του χρωμοσώματος είναι τετραγωνικός δυσδιάστατος πίνακας

5. Μέθοδος Μετάλλαξης (Mutation Operation)

Μετά από τον τελεστή διασταύρωσης πραγματοποιείται ο τελεστής μετάλλαξης. Ο τελεστής μετάλλαξης αλλάζει αυθαίρετα ένα ή περισσότερα στοιχεία, γονίδια, ενός χρωμοσώματος απογόνου. Ο ρόλος της μετάλλαξης σε ένα γενετικό αλγόριθμο είναι για να ανακτηθεί χαμένο ή αδιερεύνητο γενετικό υλικό σε ένα πληθυσμό και για να αποτρέψουν τον γενετικό αλγόριθμο να συγκλίνει σε λύση που βρίσκεται σε τοπικό ελάχιστο. [30]

Στη μέθοδο των άρθρων [24,25] χρησιμοποιείται με τυχαίο τρόπο ένας από τους εξής τρεις τελεστές μετάλλαξης: Random mutation, Non-Uniform mutation και Muhlebein's mutation.

Ας υποθέσουμε ότι το χρωμόσωμα μας έχει την πιο κάτω γενική μορφή $C(t)=(c_1, \dots, c_i, \dots, c_n)$ και ότι το στοιχείο c_i , που στην περίπτωση μας αντιστοιχεί σε μια τιμή βάρους μεταξύ δύο κόμβων είναι το γονίδιο που επιλέχθηκε για να μεταλλαχθεί. Επίσης το στοιχείο c_i ανήκει στο διάστημα $[a_i, b_i]$, όπου στην περίπτωση των τιμών των βαρών το $a_i = -1$ και το $b_i = 1$ αφού τα βάρη παίρνουν τιμές στο πεδίο $[-1,1]$.

Random Mutation [14]: Στην μέθοδο της τυχαίας μετάλλαξης, το γονίδιο c_i στο διάνυσμα χρωμοσώματος $C(t+1)=(c_1, \dots, c_i, \dots, c_n)$ είναι ένας τυχαίος αριθμός, από το πεδίο $[a_i, b_i]$. Στην περίπτωση που η κωδικοποίηση του χρωμοσώματος γίνεται με διάνυσμα δυαδικής (binary) μορφής αλλάζουν σε τυχαίες θέσεις τα γονίδια και από 0 γίνονται 1 και αντίστροφα.

Non-uniform Mutation [14, 26]: Ο τελεστής της non-uniform μετάλλαξης εφαρμόζεται σε μια γενιά t και g_{max} είναι ο μέγιστος αριθμός γενεών, τότε το γονίδιο c_i

$$C_i' = c_i + \Delta(t, b_i - c_i), \text{ if } \tau = 0 \quad (35),$$

$$C_i' = c_i + \Delta(t, c_i - a_i), \text{ if } \tau = 1 \quad (36),$$

Όπου τ είναι ένας τυχαίος αριθμός που μπορεί να έχει την τιμή 0 ή 1, και a_i, b_i είναι το κάτω και άνω όριο του πεδίου ορισμού για την μεταβλητή c_i .

Η συνάρτηση $\Delta(t, y)$ επιστρέφει μια τιμή στο διάστημα $[0,y]$ έτσι ώστε η πιθανότητα της $\Delta(t,y)$ να είναι κοντά στο 0 να αυξάνεται όσο αυξάνεται ο αριθμός των γενεών t . Λόγω αυτής της ιδιότητας ο συγκεκριμένος τελεστής εκτελεί αρχικά μια ομοιόμορφη αναζήτηση στο χώρο αναζήτησης (όταν αρχικά ο αριθμός των γενιών είναι μικρός) και σε μετέπειτα στάδια κάνει τοπική αναζήτηση. Έτσι αυξάνεται η πιθανότητα να παραχθεί νέος αριθμός ο οποίος να είναι κοντά στον διάδοχο του από το να παραχθεί μια τυχαία επιλογή. Για την συνάρτηση $\Delta(t,y)$ χρησιμοποιείται η εξής:

$$\Delta(t, y) = y \left(1 - r^{\left(1 - \frac{t}{gmax}\right)^b} \right) \quad (37),$$

Όπου r είναι ένας τυχαίος αριθμός στο διάστημα $[0..1]$, $gmax$ είναι ο μέγιστος αριθμός γενεών που μπορούν να παραχθούν και το b είναι μια παράμετρος η οποία καθορίζει το βαθμό εξάρτησης στον αριθμό της επανάληψης. Μια τιμή που θα μπορούσε να πάρει η τιμή b είναι $b=5$.

Muhlenbein's mutation [14, 17]: Η νέα τιμή του γονιδίου c_i υπολογίζεται ως εξής:

$$c'_i = c_i \pm rang_i \cdot \gamma \quad (38),$$

όπου $rang_i$ καθορίζει το εύρος μετάλλαξης και ορίζεται ως εξής, $rang_i = 0.1 \cdot (b_i - a_i)$. Το πρόσημο θετικό (+) ή αρνητικό (-) επιλέγεται με πιθανότητα 0.5 και

$$\gamma = \sum_{k=0}^{15} a_k 2^{-k} \quad (39),$$

$a_i \in \{0,1\}$ παράγεται τυχαία με πιθανότητα το a_i να είναι ίσο με 1,

$$p(a_i = 1) = \frac{1}{16}.$$

Με τη χρήση αυτού του τελεστή παράγονται τιμές στο διάστημα $[c_i - rang_i, c_i + rang_i]$. Ο συγκεκριμένος τελεστής μετάλλαξης είναι ικανός να εντοπίσει το βέλτιστο c_i με ακρίβεια ίση με $rang_i \cdot 2^{-15}$.

Στη περίπτωση κωδικοποίησης του χρωμοσώματος με τη μορφή δυσδιάστατου πίνακα χρησιμοποιούνται τρεις διαφορετικοί τελεστές μετάλλαξης, ο οριζόντιος τελεστής ο κάθετος και ο διαγώνιος όπως παρουσιάζονται πιο κάτω[19]:

Οριζόντιος τελεστής

Αρχικά επιλέγονται δύο γραμμές με τυχαίο τρόπο και εφαρμόζεται ο τελεστής στην περιοχή που έχει επιλεγεί, όπως δείχνει το σχήμα 3.3.3.4.

W11	W12	W13	W14	W15
W21	W22	W23	W24	W25
W31	W32	W33	W34	W35
W41	W42	W43	W44	W45
W51	W52	W53	W54	W55

→

W11	W12	W13	W14	W15
W25	W24	W23	W22	W21
W35	W34	W33	W32	W31
W41	W42	W43	W44	W45
W51	W52	W53	W54	W55

Σχήμα 3.3.3.4 Οριζόντιος Τελεστής μετάλλαξης όταν η κωδικοποίηση του χρωμοσώματος γίνεται με τετραγωνικό δυσδιάστατο πίνακα

Κάθετος τελεστής

Αρχικά επιλέγονται δύο στήλες με τυχαίο τρόπο και εφαρμόζεται ο τελεστής στην περιοχή που έχει επιλεγεί, όπως δείχνει το σχήμα πιο κάτω (3.3.3.5):

W11	W12	W13	W14	W15
W21	W22	W23	W24	W25
W31	W32	W33	W34	W35
W41	W42	W43	W44	W45
W51	W52	W53	W54	W55

→

W11	W12	W13	W54	W55
W21	W22	W23	W44	W45
W31	W32	W33	W34	W35
W41	W42	W43	W24	W25
W51	W52	W53	W14	W15

Σχήμα 3.3.3.5 Κάθετος Τελεστής μετάλλαξης όταν η κωδικοποίηση του χρωμοσώματος γίνεται με τετραγωνικό δυσδιάστατο πίνακα

Διαγώνιος τελεστής

Επιλέγονται δύο αριθμοί στα όρια μια στήλης και μιας γραμμής που σχηματίζουν διαγώνιο μεταξύ τους και εφαρμόζεται ο τελεστής ως εξής (σχήμα 3.3.3.6):

W11	W12	W13	W14	W15
W21	W22	W23	W24	W25
W31	W32	W33	W34	W35
W41	W42	W43	W44	W45
W51	W52	W53	W54	W55

→

W11	W45	W13	W14	W15
W21	W22	W34	W24	W25
W31	W32	W33	W23	W35
W41	W42	W43	W44	W12
W51	W52	W53	W54	W55

Σχήμα 3.3.3.6 Διαγώνιος Τελεστής μετάλλαξης όταν η κωδικοποίηση του χρωμοσώματος γίνεται με τετραγωνικό δυσδιάστατο πίνακα

6. Μέθοδος επιλογής χρωμοσωμάτων

Σύμφωνα με τον γενετικό αλγόριθμο, τα χρωμοσώματα επιλέγονται από τον πληθυσμό για να αποτελέσουν τους γονείς για να γίνει η διασταύρωση και να προκύψουν τα νέα χρωμοσώματα που είναι οι απόγονοι οι οποίοι θα αποτελέσουν τον πληθυσμό της επόμενης γενιάς. Σύμφωνα με τη θεωρία του Δαρβίνου για εξέλιξη τα καλύτερα χρωμοσώματα πρέπει να επιβιώσουν για να σχηματίζουν τα νέα χρωμοσώματα απογόνους. Υπάρχουν διάφορες μέθοδοι για να γίνει η επιλογή των χρωμοσωμάτων. [30]

Στην μέθοδο των άρθρων [24, 25] για την εκμάθηση του πίνακα συνδέσεων(βάρη) του μοντέλου των ΑΓΧ χρησιμοποιούνται η μέθοδος roulette wheel selection και η tournament selection.

Roulette wheel selection[36]: Ο συγκεκριμένος γενετικός τελεστής για επιλογή λύσεων που θα χρησιμοποιηθούν για ανασυνδυασμό γονιδίων (recombination/ crossover) είναι γνωστός και ως Fitness proportionate selection. Σε αυτή τη μέθοδο, όπως σε όλες τις μεθόδους επιλογής, η συνάρτηση καταλληλότητας (fitness function) αναθέτει μία τιμή καταλληλότητας σε όλες τις πιθανές λύσεις που αναπαριστούν τα χρωμοσώματα. Αυτή η τιμή καταλληλότητας χρησιμοποιείται για να υπολογιστεί μια πιθανότητα επιλογής η οποία συνδέεται με το κάθε χρωμόσωμα. Εάν f_i είναι η τιμή καταλληλότητας για κάθε ένα χρωμόσωμα i μέσα στον πληθυσμό τότε η πιθανότητα για να επιλεγεί υπολογίζεται ως εξής:

$$P_i = \frac{f_i}{\sum_{j=1}^N f_j} \quad (40),$$

όπου N είναι ο αριθμός όλων των χρωμοσωμάτων σε ένα πληθυσμό.

Η μέθοδος αυτή μπορούμε να φανταστούμε ότι είναι παρόμοια με τον τροχό της ρουλέτας στο καζίνο όπου γίνεται ανάθεση ενός μέρους του τροχού στις πιθανές λύσεις βάση της τιμής καταλληλότητας. Αυτό γίνεται με το να διαιρέσουμε την τιμή καταλληλότητας ενός χρωμοσώματος για επιλογή με το άθροισμα όλων των τιμών καταλληλότητας όλων των χρωμοσωμάτων του πληθυσμού, έτσι γίνεται

κανονικοποίηση στο 1. Μετά πραγματοποιείται μια τυχαία επιλογή η οποία είναι παρόμοια με τον τρόπο που γυρίζει ο τροχός της ρουλέτας.

Οι υποψήφιος λύσεις με καλύτερη τιμή καταλληλότητας έχουν μικρότερη πιθανότητα να μην επιλεγούν, αλλά παρόλα αυτά υπάρχει μια μικρή πιθανότητα να μην επιλεγούν.

Tournament Selection [37]: Η μέθοδος αυτή όπως αναφέρει και το όνομα περιλαμβάνει το τρέξιμο αρκετών «τουρνουά – tournaments» μεταξύ κάποιων χρωμοσωμάτων τα οποία επιλέγονται τυχαία από τον πληθυσμό. Ο νικητής κάθε τουρνουά, αυτός με την καλύτερη τιμή καταλληλότητας επιλέγεται για την διαδικασία του ανά-συνδυασμού (crossover). Η επιλογή ρυθμίζεται με την αλλαγή του αριθμού των τουρνουά. Ένα είναι μεγάλος ο αριθμός τότε τα αδύναμα χρωμοσώματα(ως προς την καταλληλότητα τους) έχουν μικρότερη πιθανότητα να επιλεγθούν.

Ο ψευδοκώδικας είναι ο εξής:

Επέλεξε τυχαία k (αριθμός τουρνουά) χρωμοσώματα από τον πληθυσμό

Επέλεξε το καλύτερο χρωμόσωμα από το τουρνουά με πιθανότητα p

Επέλεξε το δεύτερο καλύτερο χρωμόσωμα από το τουρνουά με πιθανότητα $p^(1-p)$*

Επέλεξε το τρίτο καλύτερο χρωμόσωμα από το τουρνουά με πιθανότητα $p^((1-p)^2)$*

κοκ.

Η ντετερμινιστική μέθοδος tournament selection επιλέγει το καλύτερο χρωμόσωμα σε κάθε τουρνουά (όταν $p=1$). Αν ο αριθμός του τουρνουά είναι $k=1$ τότε η επιλογή είναι ισοδύναμη με την τυχαία επιλογή. Τα επιλεγμένα χρωμοσώματα μπορούν να αφαιρεθούν από τον πληθυσμό, αλλιώς αυτά θα μπορούν να επιλεγθούν περισσότερα από μια φορά στην επόμενη γενεά.

7. Πιθανότητα διασταύρωσης (Probability of recombination)

Η πιθανότητα διασταύρωσης μας λέει πόσο συχνά γίνεται διασταύρωση. Εάν δεν υπάρξει διασταύρωση τα χρωμοσώματα απόγονοι είναι ακριβές αντίγραφο των γονέων

χρωμοσωμάτων. Εάν είναι 100% σημαίνει ότι όλοι οι απόγονοι σχηματίζονται από διασταύρωση ενώ αν είναι 0% όλη η νέα γενιά χρωμοσωμάτων αποτελείται από ακριβής αντίγραφα των χρωμοσωμάτων του παλιού πληθυσμού. Η διασταύρωση γίνεται με την ελπίδα ότι τα νέα χρωμοσώματα που δημιουργούνται θα έχουν καλά μέρη από τα χρωμοσώματα του παλιού πληθυσμού και ότι τα νέα θα είναι καλύτερα.[30]

Στη μέθοδο των άρθρων [24,25] προτείνεται πιθανότητα διασταύρωσης 0.9.

8. Πιθανότητα Μετάλλαξης (Probability of Mutation)

Η πιθανότητα μετάλλαξης λέει πόσο συχνά μέρη του χρωμοσώματος μεταλλάσσονται. Εάν δεν υπάρχει μετάλλαξη τότε οι απόγονοι που σχηματίζονται μετά την διαδικασία της διασταύρωσης δεν θα έχουν κάποια αλλαγή. Εάν πραγματοποιείται μετάλλαξη τότε μέρος του χρωμοσώματος αλλάζει. Όσο μεγαλύτερη πιθανότητα μετάλλαξης έχουμε τόσο μεγαλύτερο μέρος του χρωμοσώματος αλλάζει, για παράδειγμα αν η πιθανότητα είναι 100% τότε όλο το χρωμόσωμα θα αλλάξει ενώ αν είναι 0% τότε δεν θα υπάρξει αλλαγή. [30]

Στη μέθοδο των άρθρων [24,25] προτείνεται πιθανότητα μετάλλαξης 0.5.

9. Μέγεθος πληθυσμού

Μια άλλη σημαντική παράμετρος των γενετικών αλγορίθμων είναι το μέγεθος του πληθυσμού, το οποίο μας λέει πόσα χρωμοσώματα περιέχει ένας πληθυσμός (σε μία γενιά). Εάν είναι πολύ λίγα τα χρωμοσώματα ο γενετικός αλγόριθμος έχει λίγες πιθανότητες να εκτελέσει διασταύρωση και μόνο ένα μικρό μέρος του χώρου αναζήτησης εξερευνάτε. Από την άλλη εάν υπάρχουν πολλά χρωμοσώματα ο γενετικός αλγόριθμος γίνεται αργός. [30]

Στη μέθοδο των άρθρων [24,25] προτείνεται μέγεθος πληθυσμού 100 χρωμοσώματα.

10. Μέγιστος αριθμός γενεών (max_generation)

Όπως αναφέραμε πιο πάνω ένα κριτήριο τερματισμού του γενετικού αλγορίθμου είναι όταν η αναζήτηση για εύρεση λύσης στο πρόβλημα είναι ανεπιτυχής και παραχθεί ένας μέγιστος αριθμός γενεών. [24,25]

Στη μέθοδο των άρθρων [24,25] προτείνεται μέγιστος αριθμός γενεών 300 000.

11. Μέγιστο όριο τιμής καταλληλότητας (max_fitness)

Ακόμα ένα κριτήριο που αναφέραμε στον τερματισμό του γενετικού αλγορίθμου είναι αυτό της μέγιστης τιμής καταλληλότητας. Οι διάφορες υποψήφιες λύσεις αξιολογούνται βάση της συνάρτησης καταλληλότητας και οι τιμές που λαμβάνουν από αυτή τη συνάρτηση καθορίζουν πόσο κοντά βρίσκονται στην λύση. Έτσι το μέγιστο όριο καταλληλότητας είναι ένα όριο που θα κρίνει ότι η εκμάθηση είναι επιτυχημένη και έχει βρεθεί υποψήφια λύση η οποία δίνει ακολουθία καταστάσεων τιμών ενεργοποίησης ίδια ή πολύ κοντινή με αυτήν των δεδομένων εισόδου. [24, 25]

Στη μέθοδο των άρθρων [24,25] προτείνεται μέγιστο όριο τιμής καταλληλότητας 0.999.

Κεφάλαιο 4

Ανάλυση Δεδομένων

4.1 Τι είναι η Προ-εκλαμψία?	65
4.1.1 Ενδείξεις και συμπτώματα	66
4.1.2 Παράγοντες Κινδύνου	67
4.2 Δεδομένα	68
4.3 Παρουσίαση Ανάλυσης Δεδομένων	73

Στο κεφάλαιο αυτό εξηγείται το σύνδρομο της προ-εκλαμψίας, παρουσιάζονται τα δεδομένα, καθώς και οι παράγοντες που είναι κρίσιμοι για την εμφάνιση προ-εκλαμψίας. Τέλος παρουσιάζεται η ανάλυση των δεδομένων και κάποια συμπεράσματα.

4.1 Τι είναι η Προ-εκλαμψία?

Η προ-εκλαμψία (preeclampsia) είναι μια διαταραχή η οποία εκδηλώνεται μόνο κατά τη διάρκεια της εγκυμοσύνης και την περίοδο ακριβώς μετά τη γέννηση του βρέφους και επηρεάζει τόσο τη μητέρα αλλά και το έμβρυο. Η επιπλοκή αυτή επηρεάζει τουλάχιστον ένα 5 με 8% όλων των κυήσεων και εξελίσσεται πολύ γρήγορα. Χαρακτηριστικό της προ-εκλαμψίας είναι η υψηλή αρτηριακή πίεση και η παρουσία πρωτεΐνης στα ούρα. Άλλα συμπτώματα που παρουσιάζονται είναι πρήξιμο των άκρων, απότομη αύξηση βάρους, πονοκέφαλοι, διαταραχή στην όραση και πόνοι στην κοιλιά. Η προ-εκλαμψία εμφανίζεται στις εγκύους γυναίκες συνήθως μετά τις 20 εβδομάδες κυοφορίας τους (το τελευταίο 2^ο ή 3^ο τρίμηνο, ή στο μέσο προς τελευταίο στάδιο της εγκυμοσύνης), αλλά μπορεί και να εμφανιστεί νωρίτερα. Είναι πολύ σημαντικό να αναγνωριστεί και να διαχειριστεί έγκαιρα η προ-εκλαμψίας λόγω του ότι εξελίσσεται πολύ γρήγορα, πριν να εξελιχθεί σε εκλαμψία. Η εκλαμψία είναι μια πολύ επικίνδυνη κατάσταση του οργανισμού η οποία μπορεί να οδηγήσει σε θάνατο της

εγκύου γυναίκας και του εμβρύου και η οποία εκδηλώνεται με αιφνίδια εμφάνιση τονικών και κλονικών σπασμών και με απώλεια της συνείδησης (coma). Η εκλαμψία δεν είναι πολύ συνηθισμένη αφού μπορεί να παρατηρηθεί σε συχνότητα περίπου 0,2%-0,5% των κύσεων. [18, 34, 40]

Η προ-εκλαμψία είναι γνωστή και με τους όρους Pregnancy Induced Hypertension (PIH) και Preclampsia toxemia (PET). Επίσης το σύνδρομο HELLP θεωρείται μία ακόμα παραλλαγή της προ-εκλαμψίας. Εμφανίζεται και αυτό κατά τα τελευταία στάδια της εγκυμοσύνης ή μετά τη γέννηση του εμβρύου και έχει πάρει το όνομα του από τον Dr. Louis Weinstein το 1982 σύμφωνα με τα χαρακτηριστικά του τα οποία είναι: **H** (hemolysis- αιμολυσία, η διάσπαση των κόκκινων κυττάρων του αίματος), **EL** (elevated liver enzymes-αυξημένα ένζυμα στο συκώτι) και **LP** (low platelet count – μικρός αριθμός αιμοπεταλίων). Η προ-εκλαμψία καθώς και άλλες υπερτασικές διαταραχές μπορούν να αποτελέσουν αιτία για ασθένεια και θάνατο τόσο στην έγκυο γυναίκα όσο και στο ίδιο το βρέφος. Οι επιπλοκές αυτές της εγκυμοσύνης υπολογίζεται ότι είναι υπεύθυνες για 76,000 θανάτους σε μητέρες και 500,000 θανάτους σε βρέφη. [18, 34, 40]

4.1.1 Ενδείξεις και συμπτώματα

Η προ-εκλαμψία είναι αρκετά σοβαρή και πολύ επικίνδυνη επιπλοκή στην εγκυμοσύνη μιας γυναίκας επειδή πολλές φορές οι ενδείξεις και τα συμπτώματα της μοιάζουν με τα φυσιολογικά συμπτώματα εγκυμοσύνης. Η κυριότερη ένδειξη για το ότι μια έγκυος γυναίκα μπορεί να εμφανίσει προ-εκλαμψία είναι η υψηλή αρτηριακή πίεση, δηλαδή να είναι τουλάχιστον 140/90 σε δύο διαφορετικές μετρήσεις που λαμβάνονται σε διάστημα 6 ωρών ή να έχει αυξηθεί σημαντικά (πέραν των 30mm η συστολική ή πέραν των 15mm η διαστολική). Η πρωτεΐνη στα ούρα είναι ακόμα μια ένδειξη για προ-εκλαμψία. Αν η έγκυος γυναίκα σε 2 διαφορετικά δείγματα ούρων που έγιναν στο διάστημα 6 ωρών εμφανίσει τουλάχιστον 0.3 γραμμάρια πρωτεΐνης albumin σε ένα λίτρο ούρων. Επίσης οι έγκυες γυναίκες που μπορεί να εμφανίσουν προ-εκλαμψία παρουσιάζουν φούσκωμα στα άκρα τους, μπορεί φουσκώσουν τα πόδια και τα χέρια τους αλλά να παρουσιάσουν πρήξιμο στο πρόσωπο και γύρω από τα μάτια τους. Αυτό από μόνο του δεν μπορεί να θεωρηθεί σύμπτωμα για προ-εκλαμψία αφού μπορεί να συμβεί και στις έγκυες γυναίκες που έχουν φυσιολογική εγκυμοσύνη. Άλλη ένδειξη για

προ-εκλαμψία είναι η απότομη αύξηση βάρους, αν οι έγκυες γυναίκες βάζουν περισσότερο από 1 κιλό κάθε εβδομάδα. Αυτό μπορεί να συμπεριλαμβάνει και την κατακράτηση υγρών, αφού τα καταστραμμένα αγγεία αίματος επιτρέπουν περισσότερο νερό να διαρρέει και να μένει στο σώμα, χωρίς να περνά από τους νεφρούς για να φύγει. Άλλα συμπτώματα που μπορούν να συνδεθούν με την εμφάνιση της προ-εκλαμψίας είναι όταν η έγκυος γυναίκα νιώσει ξαφνική ναυτία ή να κάνει εμετούς κατά το μέσο της εγκυμοσύνης της ή όταν παραπονιέται για πόνους χαμηλά στην πλάτη. Επίσης αν η έγκυος γυναίκα νιώθει έντονους πονοκεφάλους και ημικρανίες που δεν λένε να περάσουν είναι λόγος ανησυχίας για τυχόν εμφάνιση προ-εκλαμψίας. Ένα άλλο σοβαρό σύμπτωμα είναι όταν η έγκυος γυναίκα αισθανθεί αλλαγές στην όραση της, όπως προσωρινή απώλεια της όρασης, να αισθάνεται φώτα που αναβοσβήνουν, να βλέπει θολά ή να βλέπει κηλίδες. Οι αλλαγές αυτές στην όραση μπορεί να συνδέονται με ενόχληση του κεντρικού νευρικού συστήματος ή φούσκωμα του εγκεφάλου. Σε σοβαρές περιπτώσεις προ-εκλαμψίας, η αύξηση της αρτηριακής πίεσης και η αύξηση της ποσότητας πρωτεΐνης στα ούρα είναι πολύ πιο έντονη. [18, 34]

4.1.2 Παράγοντες Κινδύνου

Η προ-εκλαμψία καλείται η ‘ασθένεια των υποθέσεων’ αφού τα ακριβή αίτια που την προκαλούν δεν είναι γνωστά αλλά ο εντοπισμός της βασίζεται σε ενδείξεις και υποθέσεις.[34] Οι επιστήμονες έχουν καταλήξει σε κάποια συμπεράσματα βάσει στατιστικών μελετών για τους παράγοντες που αυξάνουν την πιθανότητα μια έγκυος γυναίκα να αναπτύξει προ-εκλαμψία. Γυναίκες με Afro-Caribbean εθνικότητα, δηλαδή γυναίκες μαύρης φυλής έχουν περισσότερες πιθανότητες για εμφάνιση προ-εκλαμψίας.[4,5] Επίσης περισσότερες πιθανότητες έχουν οι γυναίκες που μένουν έγκυες για πρώτη φορά. Ακόμα η εγκυμοσύνη σε μικρή ηλικία και κυρίως κατά τη διάρκεια της εφηβείας (13-18 χρονών) ή σε ηλικία από 37 ετών και άνω αυξάνουν τις πιθανότητες για εμφάνιση της προ-εκλαμψίας. Οι ηλικίες με τα μικρότερα ποσοστά εμφάνισης της πάθησης είναι 19-25 ετών. Όσο μεγαλύτερη είναι η απόκλιση της ηλικίας από αυτά τα όρια τόσο μεγαλώνει η συχνότητα εμφάνισης της προ-εκλαμψίας. Οι γυναίκες που κυοφορούν περισσότερα από ένα έμβρυα διατρέχουν επίσης μεγαλύτερες πιθανότητες εμφάνισης της προ-εκλαμψίας. Οι υπέρβαρες γυναίκες, με υψηλή ένδειξη δείκτη μάζας σώματος (body-mass index, bmi), οι γυναίκες που

πάσχουν από διαβήτη, οι γυναίκες που είχαν προηγούμενο ιστορικό προ-εκλαμψίας στην οικογένεια τους είτε εμφάνισε η μητέρα ή η αδερφή τους, όλες αυτές έχουν περισσότερες πιθανότητες να εμφανίσουν προ-εκλαμψία. Ακόμα οι γυναίκες που πάσχουν από άσθμα θα πρέπει να προσέχουν ιδιαίτερα λόγω του ότι μπορεί να προκληθούν παροξυσμοί της νόσου στη μήτρα, με αποτέλεσμα την περιορισμένη παροχή οξυγόνου στο έμβρυο. Αυτό μπορεί να οδηγήσει σε επιπλοκές, μεταξύ των οποίων η υπέρταση και η προ-εκλαμψία. Επίσης οι έγκυες που πάσχουν από σιδηροπενική αναιμία διατρέχουν διπλάσιο κίνδυνο να εμφανίσουν προ-εκλαμψία αλλά και οι γυναίκες που πριν από την έναρξη της εγκυμοσύνης τους «υπέφεραν» από υπέρταση, έχουν μεγαλύτερο κίνδυνο να εμφανίσουν προ-εκλαμψία κατά την κύηση. [39]

Η προ-εκλαμψία είναι μια πολύ σοβαρή επιπλοκή στην εγκυμοσύνη και θέτει σε κίνδυνο τη ζωή και την υγεία τόσο της ίδιας της μητέρας αλλά και του ίδιου του εμβρύου. Έτσι είναι σημαντικό να γίνει έγκαιρα η διάγνωση της. Βέβαια αυτό είναι αρκετά δύσκολο και η πρόγνωση της ακόμα πιο δύσκολη. Για τους πιο πάνω λόγους οποιοδήποτε εργαλείο το οποίο θα προβλέπει την εμφάνιση της προ-εκλαμψίας, ή μια υπολογιστική μέθοδος για την έγκαιρη αναγνώριση των περιπτώσεων υψηλού κινδύνου, θα ήταν πολύ βοηθητικό για τους γυναικολόγους – μαιευτήρες αλλά και για τις εγκύους γυναίκες. Στην βιβλιογραφία υπάρχουν προηγούμενες μελέτες στις οποίες χρησιμοποιήθηκαν νευρωνικά δίκτυα και άλλα έξυπνα υπολογιστικά συστήματα που προβλέπουν την εμφάνιση της προ-εκλαμψίας. [4,5]

4.2 Δεδομένα

Τα δεδομένα που είχαμε στη διάθεση μας χρησιμοποιήθηκαν σε μελέτες για κατασκευή νευρωνικών δικτύων για πρόβλεψη εμφάνισης της προ-εκλαμψίας. Αποτελούνται από περιπτώσεις γυναικών που ζουν στο Ηνωμένο Βασίλειο και παραχωρήθηκαν από το Ερευνητικό κέντρο Harris Birthright για εμβρυϊκή ιατρική (Harris Birthright Research Centre for Fetal Medicine) στο Λονδίνο. [4,5]

Όλες οι γυναίκες επισκέπτονταν το κλινικό κέντρο για τον έλεγχο ρουτίνας ώστε να διαπιστωθεί αν διατρέχουν κίνδυνο για χρωμοσωμικές ανωμαλίες. Για τον έλεγχο αυτό γίνονται διάφορες μετρήσεις μεταξύ της 11^{ης} με 13^{ης}+ 6 (μέρες), εβδομάδας κύησης.

Γίνεται το υπερηχογράφημα αυχενικής διαφάνειας στο οποίο μετριέται το πάχος του υγρού στον αυχένα του εμβρύου (fetal nuchal translucency thickness) και αναλύσεις του αίματος της μητέρας για προσδιορισμό της ελεύθερης β-χοριακής γοναδοτροπίνης (free human chorionic gonadotropin) και της πρωτεΐνης PAPP-a (pregnancy-associated plasma protein A). Επίσης μετριέται το μήκος του εμβρύου από την κορυφή μέχρι τους γλουτούς (crown rump length) και εξάγεται η ηλικία κύησης (gestational age), δηλαδή η ηλικία του εμβρύου η οποία μετριέται από την τελευταία έμμηνο ρύση της μητέρας ή 14 ημέρες πριν την σύλληψη. [4,5]

Οι γυναίκες οι οποίες συμφώνησαν να λάβουν μέρος στην μελέτη, η οποία εγκρίθηκε από την επιτροπή ηθικής του νοσοκομείου στο King's College έδωσαν γραπτώς την συγκατάθεση τους. Ζητήθηκε από τις ασθενείς να συμπληρώσουν ένα ερωτηματολόγιο στο οποίο ζητούσαν πληροφορίες όπως την ηλικία της εγκύου μητέρας, την εθνική καταγωγή της (Λευκή, αфро-καραϊβική, Ινδή, Πακιστανή, Κινέζικη, Ιαπωνική ή ανάμικτη), αν κάπνιζε τσιγάρο ή κατανάλωνε αλκοόλ ή έκανε χρήση ναρκωτικών ουσιών κατά τη διάρκεια της εγκυμοσύνης, το ιατρικό ιστορικό της, τυχόν ιατρική φαρμακευτική αγωγή, πόσες φορές γέννησε στο παρελθόν, αν είχε προηγούμενη εγκυμοσύνη με προ-εκλαμψία, η αν εμφάνισε μέλος της οικογένειας ιστορικό για preeclampsia (αδερφή, μητέρα ή και οι δύο να είχαν preeclampsia).

Επίσης στοιχεία όπως το βάρος και το ύψος της μητέρας μετρήθηκαν και υπολογίστηκε ο δείκτης μάζας σώματος (bmi) σε χιλιόγραμμα προς τετραγωνικά μέτρα. Η αρτηριακή πίεση (blood pressure-BP) των εγκύων μητέρων μετρήθηκε από αυτόματες συσκευές και στα δεδομένα χρησιμοποιήθηκε η υψηλότερη τελική μέση αρτηριακή πίεση.

Για την μέτρηση του uterine artery pulsatility index (δείκτης παλμού αρτηρίας της μήτρας) γίνεται μια τοξοειδής τομή για κάθε αρτηρία του τραχήλου της μήτρας. Χρησιμοποιήθηκαν παλμικά κύματα Doppler με θύρα δειγματοληψίας ρυθμισμένη στα 2mm για την κάλυψη όλου του αγγείου, γωνία του υπέρηχου μικρότερη από 50ο. Η μέτρηση του uterine pulsatility index (UPI) έγινε αφού λήφθηκαν τρεις όμοιες και διαδοχικές κυματομορφές και υπολογίστηκε το μέσο pulsatility index (PI) των αριστερών και των δεξιών αρτηριών.

Στο Παράρτημα Β παρουσιάζονται όλοι οι παράγοντες για τους οποίους λήφθηκαν μετρήσεις για τις περιπτώσεις γυναικών σε ένα από τα αρχεία που είχαμε. Ένας από τους παράγοντες είναι το outcome, δηλαδή το αποτέλεσμα της εγκυμοσύνης σε κάθε μια από τις περιπτώσεις και αν εμφάνισε τελικά ή όχι προ-εκλαμψία η έγκυος γυναίκα.

Από τη βιβλιογραφία που μελετήθηκε και βασισμένοι σε προηγούμενες έρευνες [4,5] κάποιοι παράγοντες φαίνεται να επηρεάζουν περισσότερο για την εμφάνιση προ-εκλαμψίας. Για αυτό το λόγο έγινε περαιτέρω ανάλυση στους παράγοντες που θεωρούνται πιο κρίσιμοι για την εμφάνιση προ-εκλαμψίας σε μια έγκυο γυναίκα. Το αρχείο δεδομένων που χρησιμοποιήθηκε ήταν διαφορετικό από το αρχείο του οποίου τους παράγοντες ανέφερα πιο πάνω. Στον πίνακα 4.2.2 παρουσιάζονται οι παράγοντες αυτοί καθώς και μια μικρή διευκρίνιση τους.

Όνομα Παράγοντα	Περιγραφή
Output	Το αποτέλεσμα για εμφάνιση ή όχι προ-εκλαμψίας της εγκύου γυναίκας. Οι τιμές του παράγοντα αυτού έχουν ως ακολούθως: Normal – Η έγκυος γυναίκα δεν εμφάνισε προ-εκλαμψία PE early <34w – Η έγκυος γυναίκα εμφάνισε προ-εκλαμψία, σε πρώιμο στάδιο της εγκυμοσύνης της, πριν τις 34 εβδομάδες κύησης. PE middle 34-37w - Η έγκυος γυναίκα εμφάνισε προ-εκλαμψία, σε μέσο στάδιο της εγκυμοσύνης της, μεταξύ 34-37 εβδομάδων κύησης PE late 38-42w - Η έγκυος γυναίκα εμφάνισε προ-εκλαμψία, σε προχωρημένο στάδιο της εγκυμοσύνης της μεταξύ 38-42 εβδομάδων κύησης
MA	Ηλικία της εγκύου μητέρας
Race	Φυλή της εγκύου μητέρας. Οι τιμές του παράγοντα αυτού έχουν ως ακολούθως: Black East Asian South Asian

	White Mixed
Height	Το ύψος της εγκύου μητέρας
Weight	Το βάρος της εγκύου μητέρας
Conception	Ο τρόπος με τον οποίο έμεινε έγκυος η γυναίκα. Οι τιμές του παράγοντα αυτού έχουν ως ακολούθως: Spontaneous: Αυθόρμητα με φυσιολογική σεξουαλική επαφή Ovulation Drugs: Όταν η έγκυος γυναίκα παίρνει φάρμακα για να προκληθεί ωορρηξία.
Smoking	Ο παράγοντας αυτός αναφέρει αν η έγκυος γυναίκα κάπνιζε ή όχι
Previous Pregnancies	Ο παράγοντας αυτός αναφέρει το ιστορικό προηγούμενων εγκυμοσύνων της εγκύου γυναίκας. Οι τιμές του παράγοντα αυτού έχουν ως ακολούθως: No previous Pregnancies – Δεν είχε προηγούμενες εγκυμοσύνες (είναι η πρώτη της εγκυμοσύνη) Previous Pregnancies with PE – Είχε προηγούμενες εγκυμοσύνες στις οποίες εμφάνισε προ-εκλαμψία Previous Pregnancies without PE –Είχε προηγούμενες εγκυμοσύνες στις οποίες δεν είχε εμφανίσει προ-εκλαμψία.
Patients mother had PE	Ο παράγοντας αυτός αναφέρει κατά πόσο η μητέρα της εγκύου γυναίκας εμφάνισε προ-εκλαμψία σε προηγούμενη της εγκυμοσύνη. Οι τιμές του παράγοντα έχουν ως ακολούθως: -No – Δεν εμφάνισε προ-εκλαμψία -Patient's mother had PE – Εμφάνισε προ-εκλαμψία
Chronic hypertension	Ο παράγοντας αυτός αναφέρει κατά πόσο η έγκυος μητέρα πάσχει από χρόνια υπέρταση. Οι τιμές του παράγοντα έχουν ως ακολούθως:

	-No – Δεν υπέφερε από χρόνια υπέρταση -Chronic hypertension – Υπέφερε από χρόνια υπέρταση
Blood test PAPPA	Ο παράγοντας αυτός είναι η τιμή της πρωτεΐνης PAPP-A, από τις αναλύσεις αίματος της εγκύου μητέρας. Πρωτεΐνη που παράγεται από τον πλακούντα και το τοίχωμα της μήτρας κατά τη διάρκεια της εγκυμοσύνης. Μετρείται σε μονάδες ανά λίτρο (Units per Litre).
Blood test β-hCG	Ο παράγοντας αυτός είναι η τιμή της ορμόνης Beta human chorionic gonadotropin, από τις αναλύσεις αίματος της εγκύου μητέρας. Είναι μια ορμόνη που παράγεται κατά την ανάπτυξη της εγκυμοσύνης και στη συνέχεια παράγεται σε μεγάλες ποσότητες από τον πλακούντα.
Doppler	Ο παράγοντας αυτός είναι η τιμή του Doppler test (uterine pulsatility index-UPI) Υπολογίζεται από τις κυματομορφές του υπερηχογραφήματος Doppler. Είναι ένας δείκτης της ταχύτητας της ροής του αίματος σε αρτηρίες της μήτρας. [31]
Blood pressure	Ο παράγοντας αυτός αναφέρεται στην αρτηριακή πίεση της εγκύου γυναίκας

Πίνακας 4.2.2 Πίνακας πιο κρίσιμων παραγόντων

Έχοντας δεδομένο το αρχείο με τις παραμέτρους του πιο πάνω πίνακα και τις πραγματικές τιμές για αυτές τις παραμέτρους από διάφορες περιπτώσεις εγκύων γυναικών έγινε μια προσπάθεια για εξαγωγή κάποιων συμπερασμάτων. Τα συμπεράσματα τα οποία θα εξάγονταν θα βοηθήσουν στο επόμενο στάδιο για ανάπτυξη ενός μοντέλου AGX το οποίο θα έχει ως απώτερο σκοπό να προβλέψει εάν μια έγκυος γυναίκα θα εμφανίσει προ-εκλαμψία.

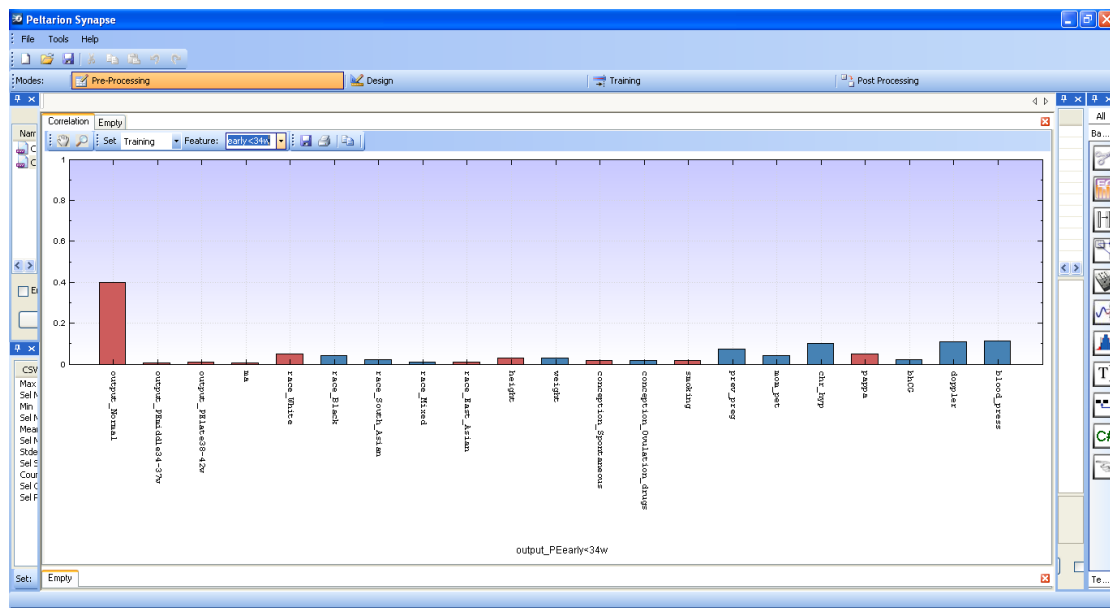
4.3 Παρουσίαση Ανάλυσης Δεδομένων

Για την ανάλυση αυτή των δεδομένων έγινε χρήση του εργαλείου Peltarion Synapse. Το εργαλείο αυτό αποτελεί ένα περιβάλλον ανάπτυξης νευρωνικών δικτύων και προσαρμοσμένων συστημάτων παρέχοντας στον χρήστη διάφορα χαρακτηριστικά. Το Peltarion Synapse προσφέρει στους χρήστες ένα περιβάλλον διεπαφής μαζί με εργαλεία και αλγορίθμους έτσι ώστε να μπορέσουν οι ίδιοι να αναλύσουν και να επεξεργαστούν τα δεδομένα τους αλλά και να αναπτύξουν ένα σύστημα το οποίο να εκπαιδεύεται με τα ίδια τα δεδομένα για την επίλυση κάποιου προβλήματος. [33]

Για την ανάλυση των δεδομένων με το εργαλείο Peltarion Synapse έγινε μια μετατροπή του αρχείου δεδομένων στην Excel σε ένα αρχείο κειμένου (text file) όπου οι τιμές των δεδομένων χωρίζονταν με απλό διάστημα (space). Το σύνολο των δεδομένων του οποίου έγινε ανάλυση στο εργαλείο αποτελείται από 11508 περιπτώσεις εγκύων γυναικών από τις οποίες οι 11103 έχουν τιμή Normal στην παράμετρο output ενώ οι υπόλοιπες 405 περιπτώσεις έχουν τιμή Pet Early <34weeks, ή Pet Middle 34-37 weeks, ή Pet Late 38-42 weeks.

Το εργαλείο μετά την εισαγωγή του αρχείου των δεδομένων δίνει επιλογή για απεικόνιση τους σε ποικιλία γραφικών παραστάσεων. Πιο κάτω παραθέτω κάποιες γραφικές παραστάσεις από την εξερεύνηση που έγινε στα δεδομένα καθώς και τα συμπεράσμα που παίρνουμε από αυτές.

Στις παραστάσεις γραμμικών συσχετίσεων ενός παράγοντα σε σχέση με τους υπόλοιπους παράγοντες, μπορούμε να δούμε μια γενική εικόνα για όλους τους παράγοντες και σε ποιο βαθμό συσχετίζονται μεταξύ τους. Οι απεικονίσεις αυτές δείχνουν μια γραμμική συσχέτιση μεταξύ του παράγοντα που εξετάσουμε και φαίνεται στον άξονα x με τους υπόλοιπους παράγοντες. Οι μπλε ράβδοι υποδεικνύουν θετική συσχέτιση ενώ οι κόκκινες ράβδοι υποδεικνύουν αρνητική συσχέτιση. Όσο πιο ψηλά είναι η ράβδος τόσο πιο πολύ συσχετίζονται οι παράγοντες μεταξύ τους.

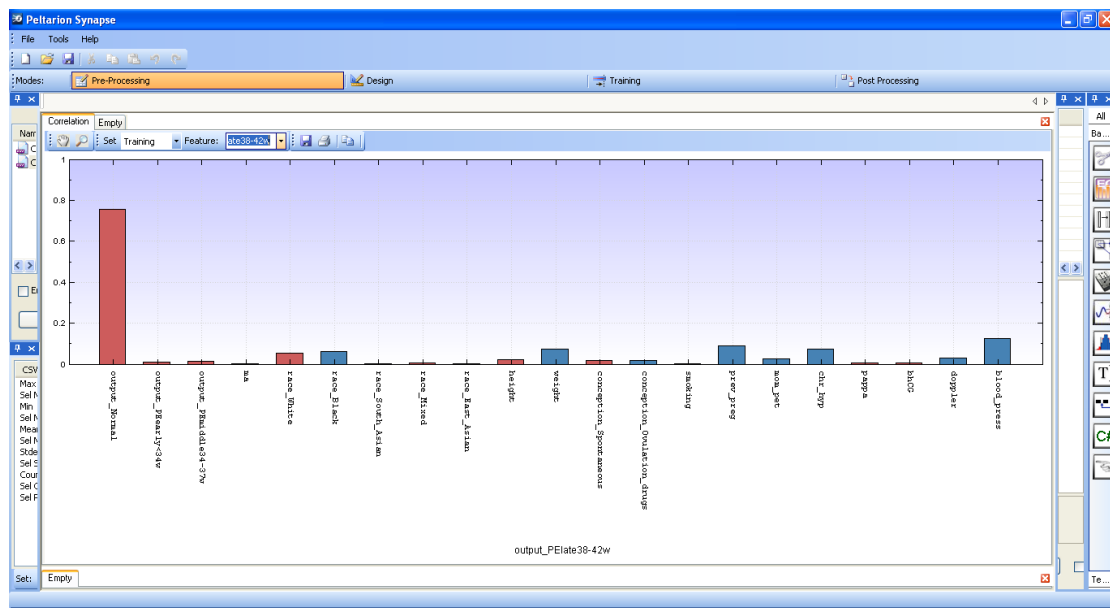


Σχήμα 4.3.1 Γραμμική συσχέτιση παράγοντα output PE early <34w με υπόλοιπους.

Η γραφική του σχήματος 4.3.1 απεικονίζει τη γραμμική συσχέτιση του παράγοντα output PE early <34w σε σχέση με τους υπόλοιπους παράγοντες. Το αποτέλεσμα του output PE early <34w είναι 1 όταν η έγκυος γυναίκα εμφάνισε προ-εκλαμψία σε πρώιμο στάδιο της εγκυμοσύνης της, γι αυτό βλέπουμε την αρνητική συσχέτιση μεταξύ του output normal, output PE middle 34-37w και output late 38-42w αφού έχουν τιμή 0. Παρατηρείται θετική συσχέτιση με τους παράγοντες blood pressure, Doppler, chronic hypertension, previous pregnancy, mom pet και race black, ενώ παρατηρείται πιο μικρή θετική συσχέτιση με τους παράγοντες bhCG, conception ovulation drugs, weight, race south Asian και race mixed. Η θετική αυτή συσχέτιση που υπάρχει μεταξύ κάποιων παραγόντων επιβεβαιώνεται και από την βιβλιογραφία. Δηλαδή γυναίκες που έχουν υψηλή αρτηριακή, πάσχουν από χρόνια υπέρταση, εμφάνισαν προ-εκλαμψία σε προηγούμενη τους εγκυμοσύνη, η μητέρα τους εμφάνισε προ-εκλαμψία στην εγκυμοσύνη της ή ανήκουν στην μαύρη φυλή έχουν περισσότερες πιθανότητες να εμφανίσουν προ-εκλαμψία. Η θετική συσχέτιση του παράγοντα output PE early <34w με τον παράγοντα Doppler υπονοεί ότι μια έγκυος γυναίκα που εμφάνισε προ-εκλαμψία σε πρώιμα στάδια της εγκυμοσύνης της έχει υψηλή τιμή στο αποτέλεσμα της εξέτασης του Doppler. Ακόμα παρατηρούμε μικρή θετική συσχέτιση με τον παράγοντα conception ovulation drugs που από αυτό συμπεραίνουμε ότι οι γυναίκες που εμφάνισαν προ-εκλαμψία χρησιμοποίησαν φάρμακα για να γίνει παραγωγή ωαρίων και να μείνουν έγκυες. Παρατηρούμε επίσης ότι η συσχέτιση του παράγοντα output PE

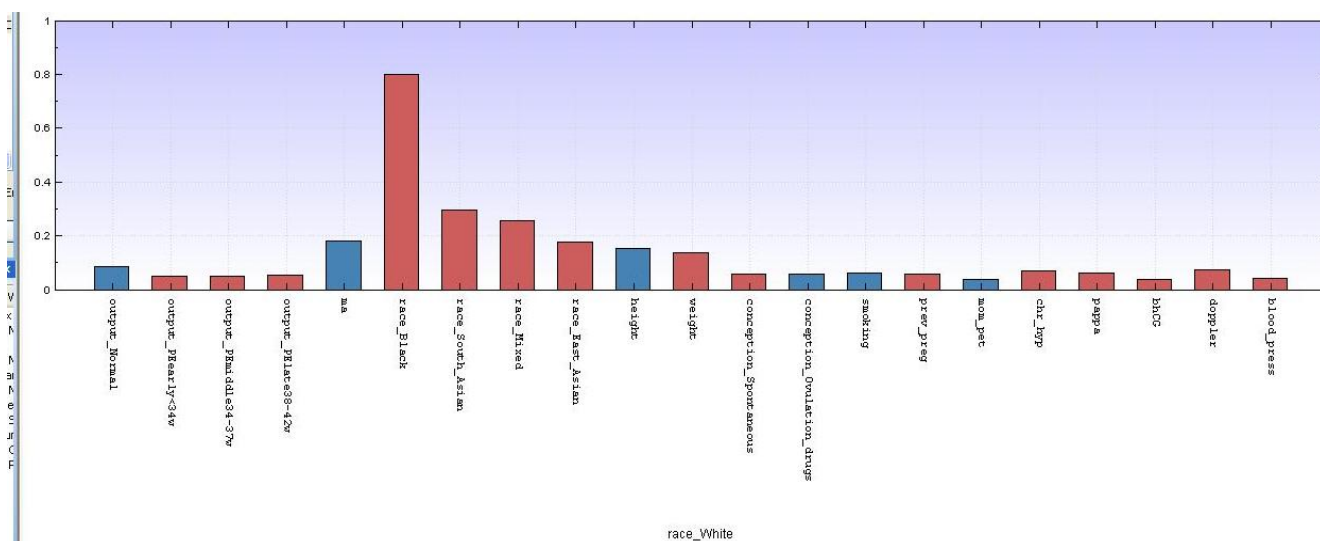
Figure 1: Correlation plot of the output variable with the input variables. The plot shows the correlation coefficient for each input variable. The output variable is 'output_PEMiddle34-37w'. The input variables are: output_Bornal, output_Mean34-37w, output_P1463B-40w, age, race_White, race_Black, race_Indian, race_Mixed, race_East_African, height, weight, occupation_Sportswomen, occupation_Divulcation, driving, smoking, previous_preg, non_preg, dm_type, papppa, hmo, outputar, and blood_press. The correlation coefficients are: output_Bornal (0.5), output_Mean34-37w (0.01), output_P1463B-40w (0.02), age (0.01), race_White (0.05), race_Black (0.03), race_Indian (0.02), race_Mixed (0.01), race_East_African (0.01), height (0.03), weight (0.03), occupation_Sportswomen (0.01), occupation_Divulcation (0.01), driving (0.01), smoking (0.01), previous_preg (0.05), non_preg (0.01), dm_type (0.05), papppa (0.03), hmo (0.01), outputar (0.05), and blood_press (0.05).

Η γραφική στο σχήμα 4.3.2 απεικονίζει τη γραμμική συσχέτιση του παράγοντα output PE middle 34-37w σε σχέση με τους υπόλοιπους παράγοντες. Το αποτέλεσμα του output PE middle 34-37w είναι 1 όταν η έγκυος γυναίκα εμφάνισε προ-εκλαμψία σε μέσο στάδιο της εγκυμοσύνης της, γι αυτό βλέπουμε την αρνητική συσχέτιση μεταξύ του output normal, output PE early <34w και output late 38-42w αφού έχουν τιμή 0. Παρατηρούμε ότι υπάρχει θετική συσχέτιση με τους ίδιους παράγοντες της προηγούμενης γραφικής, και αυτό αναμέναμε να δούμε σύμφωνα και με την βιβλιογραφία για τους παράγοντες που επηρεάζουν στην εμφάνιση προ-εκλαμψίας. Επιπλέον παρατηρούμε μικρή συσχέτιση με τον παράγοντα της ηλικίας της εγκύου μητέρας και πάρα πολύ μικρή με τον παράγοντα του καπνίσματος. Από αυτό συμπεραίνουμε ότι έγκυες γυναίκες μεγάλης ηλικίας μπορεί να παρουσιάσουν προ-εκλαμψία στο μέσο της εγκυμοσύνης τους και ότι έγκυες γυναίκες που κάπνιζαν κατά την εγκυμοσύνη τους εμφάνισαν προ-εκλαμψία σε μέσο στάδιο.

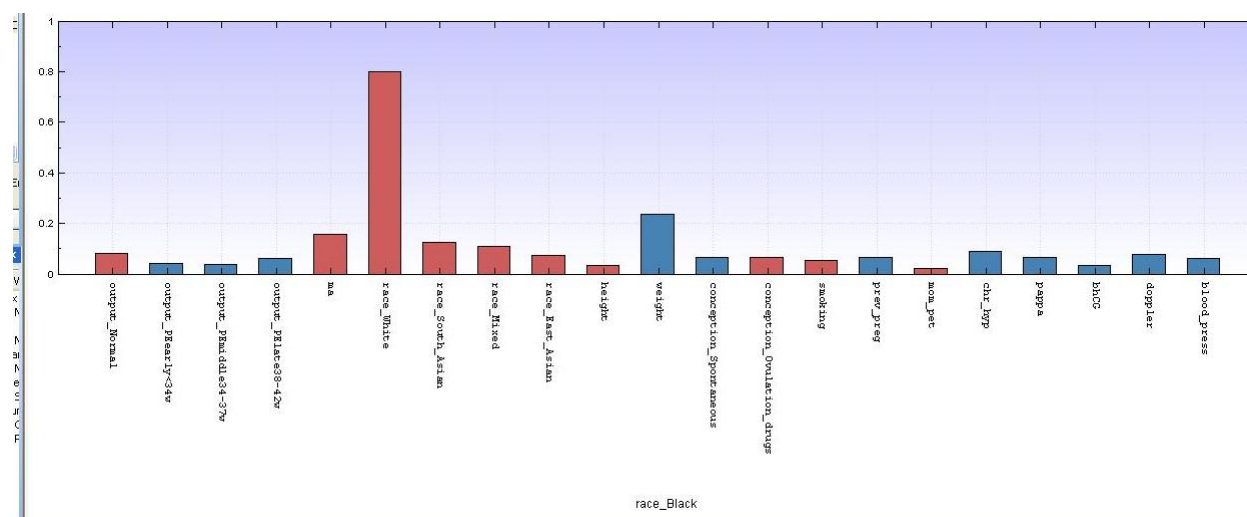


Σχήμα 4.3.3 Γραμμική συσχέτιση παράγοντα output PE late 38-42w με υπόλοιπους.

Η γραφική του σχήματος 4.3.3 απεικονίζει τη γραμμική συσχέτιση του παράγοντα output PE late 38-42w σε σχέση με τους υπόλοιπους παράγοντες. Το αποτέλεσμα του output PE middle 34-37w είναι 1 όταν η έγκυος γυναίκα εμφάνισε προ-εκλαμψία σε προχωρημένο στάδιο της εγκυμοσύνης της, γι αυτό βλέπουμε την αρνητική συσχέτιση μεταξύ του output normal, output PE early <34w και output middle 34-37w αφού έχουν τιμή 0. Παρατηρούμε ότι και εδώ υπάρχει θετική συσχέτιση με τους ίδιους παράγοντες της προηγούμενης γραφικής.



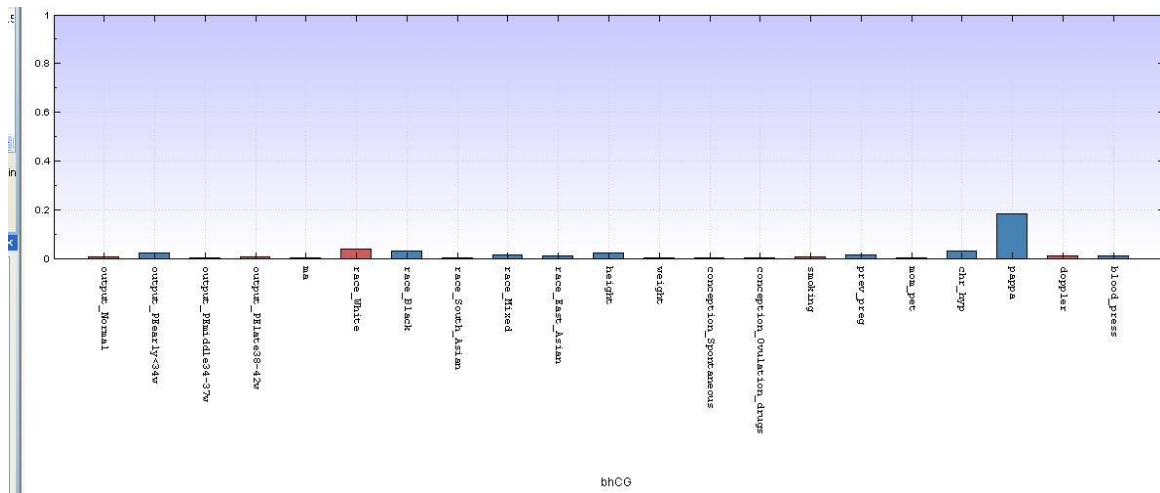
Σχήμα 4.3.4 Γραμμική συσχέτιση παράγοντα race White με υπόλοιπους



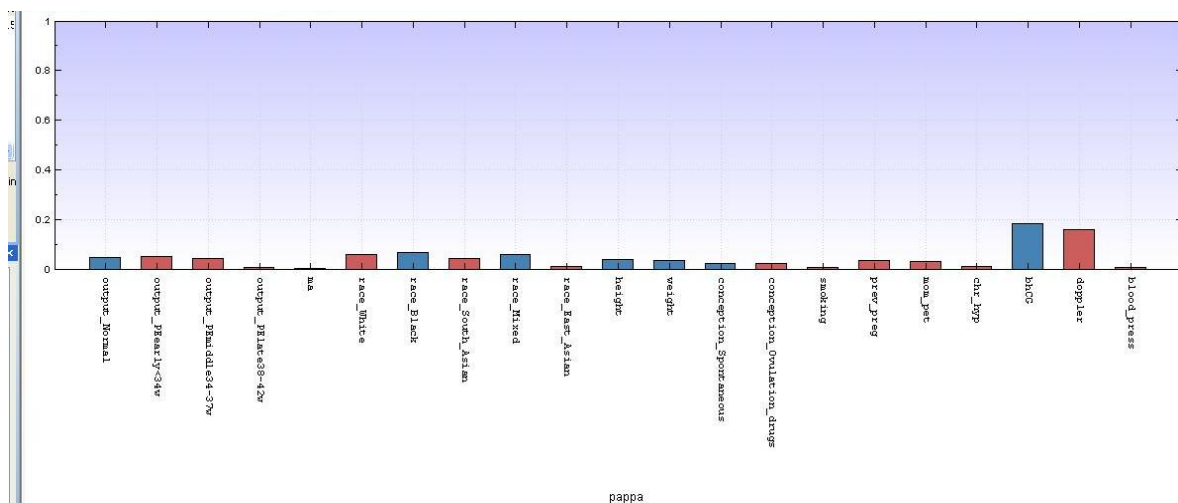
Σχήμα 4.3.5 Γραμμική συσχέτιση παράγοντα race Black (κάτω) με υπόλοιπους

Οι γραφικές στα σχήματα 4.3.4 και 4.3.5 απεικονίζουν την γραμμική συσχέτιση των παραγόντων race White και race Black σε σχέση με τους υπόλοιπους παράγοντες. Αυτό που παρατηρείται είναι ότι οι γυναίκες που ανήκουν στη μαύρη φυλή έχουν θετική συσχέτιση με τους παράγοντες output PE early <34w, output PE middle 34-37w και output late 38-42w σε αντίθεση με τις γυναίκες της λεύκης φυλής που έχουν αρνητική συσχέτιση με αυτές τις μεταβλητές. Αυτό είναι κάτι που επιβεβαιώνεται και από τη θεωρία αφού οι γυναίκες μαύρης φυλής έχουν μεγαλύτερες πιθανότητες για

εμφάνιση προ-εκλαμψίας. Ακόμα βλέπουμε έντονη συσχέτιση του παράγοντα μαύρη φυλή με το βάρος, που υπονοεί ότι οι γυναίκες μαύρης φυλής έχουν μεγαλύτερο βάρος από γυναίκες άλλων φυλών και όπως αναφέραμε οι υπέρβαρες γυναίκες μπορεί να εμφανίσουν προ-εκλαμψία. Επίσης παρατηρείται θετική συσχέτιση και με άλλους παράγοντες κινδύνου για εμφάνιση προ-εκλαμψίας, δηλαδή της αρτηριακής πίεσης, της χρόνιας υπέρτασης και προηγούμενης εγκυμοσύνης με εμφάνιση προ-εκλαμψίας.

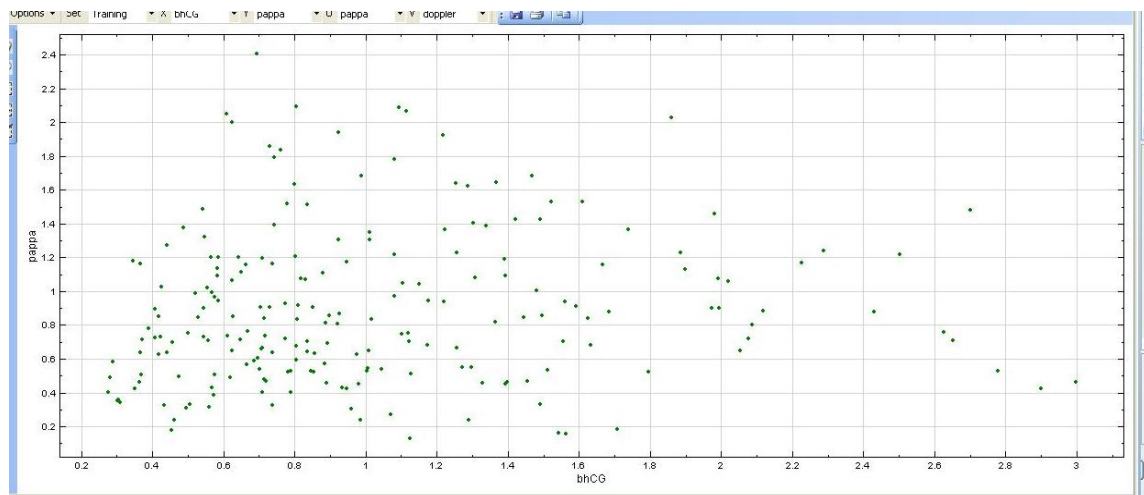


Σχήμα 4.3.6 Γραμμική συσχέτιση παράγοντα bhCG σχέση με τους υπόλοιπους.



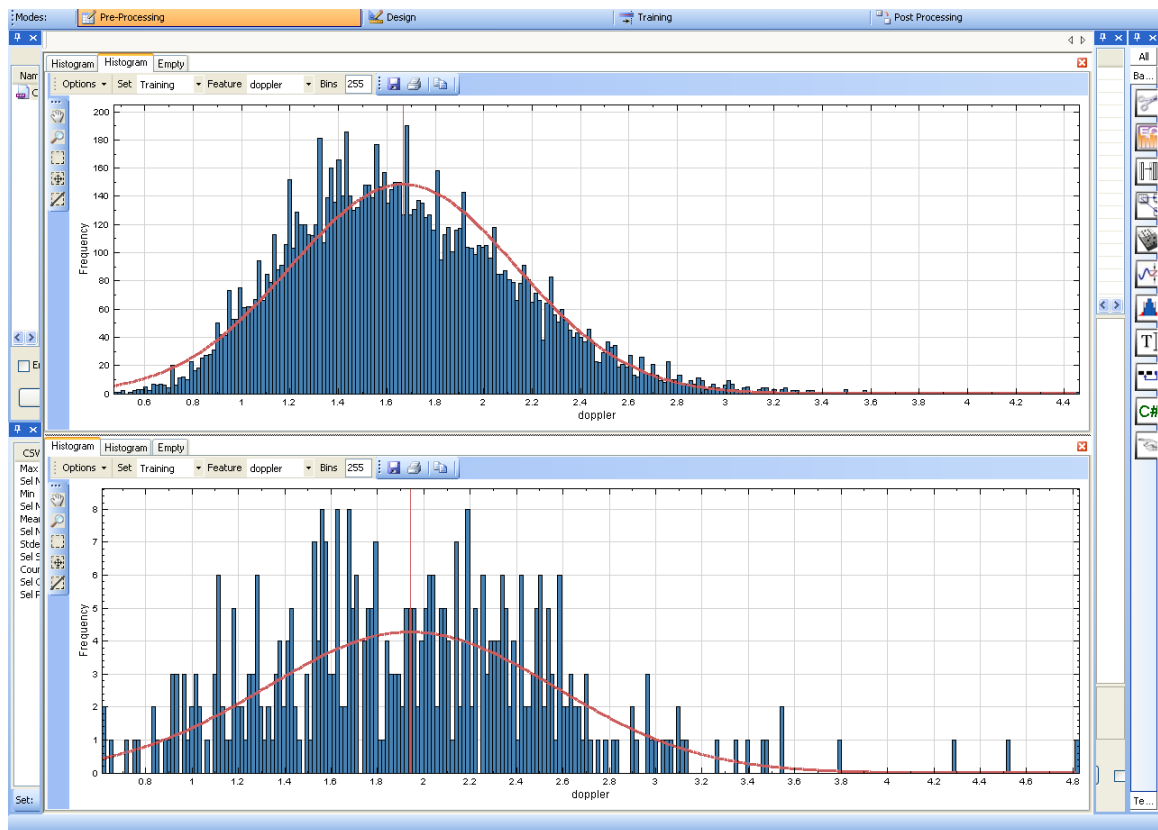
Σχήμα 4.3.7 Γραμμική συσχέτιση παράγοντα PAPP σε σχέση με τους υπόλοιπους

Οι γραφικές στα σχήματα 4.3.6 και 4.3.7 απεικονίζουν την γραμμική συσχέτιση των παραγόντων bhCG και pappA σε σχέση με τους υπόλοιπους παράγοντες. Αυτό που παρατηρείται έντονα είναι η θετική συσχέτιση του παράγοντα bhCG με παράγοντα pappA.



Σχήμα 4.3.8 Καρτεσιανή συσχέτιση παράγοντα Pappa με bhCG για το σύνολο εγκύων γυναικών που εμφάνισαν προ-εκλαμψία.

Με την γραφική παράσταση του σχήματος 4.3.8 η οποία αποτελεί ένα καρτεσιανό διάγραμμα στο οποίο γίνεται προβολή των δύο παραγόντων για το σύνολο των εγκύων γυναικών που εμφάνισαν προ-εκλαμψία προσπαθήσαμε να βρούμε μια σχέση μεταξύ bhCG και pappa. Στο σύνολο αυτό έγινε χρήση του εργαλείου το οποίο αφαιρεί κάποια δεδομένα τα οποία μπορεί να είναι εκτός του κανονικού πεδίου τιμών. Στον άξονα τον x αναπαριστάτε η bhCG ενώ στον άξονα τον y αναπαριστάτε η pappa. Αυτό που μπορούμε να δούμε από την γραφική είναι ότι σε κάποια σημεία όταν αυξάνονται τα επίπεδα τιμών της bhCG, αυξάνονται και τα επίπεδα της pappa. Αλλά δεν μπορούμε να το πούμε απόλυτα επειδή είναι αρκετά ομοιόμορφα κατανεμημένα στον χώρο.

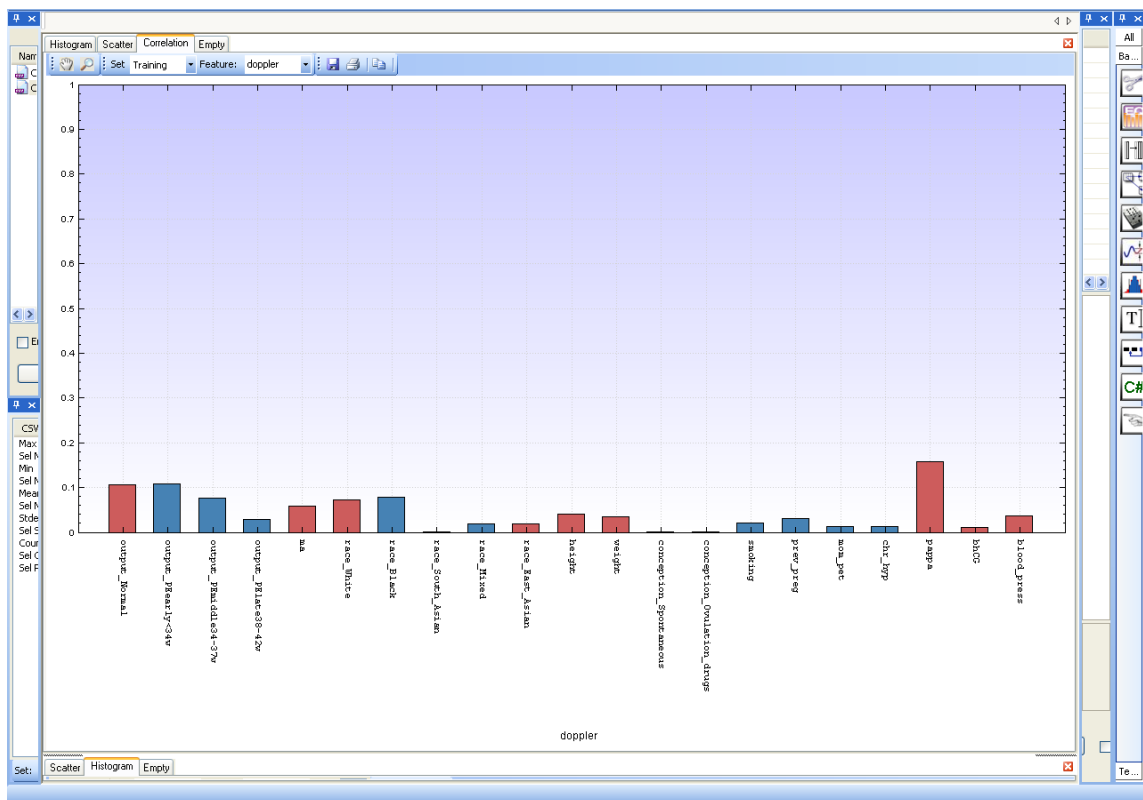


Σχήμα 4.3.9 Ιστογράμματα τιμών παράγοντα Doppler στις περιπτώσεις που δεν εμφάνισαν προ-εκλαμψία (πρώτο ιστόγραμμα) και στις περιπτώσεις που εμφάνισαν (δεύτερο ιστόγραμμα).

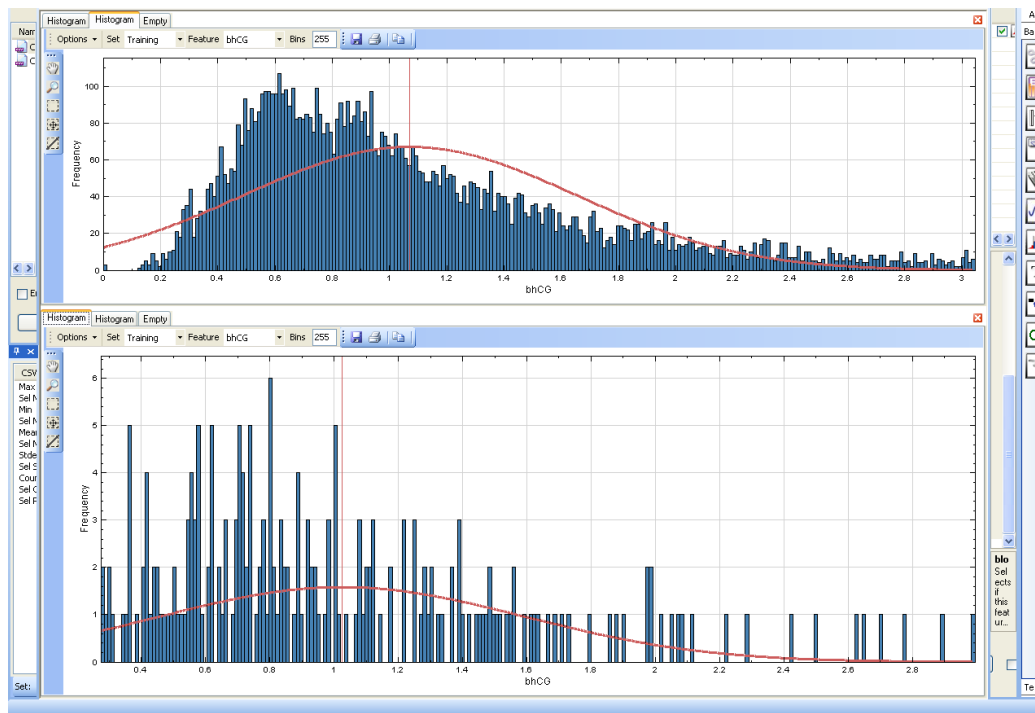
Στις γραφικές στο σχήμα 4.3.9 παρουσιάζονται τα ιστογράμματα του αποτελέσματος του Doppler test στις περιπτώσεις των εγκύων γυναικών που δεν εμφάνισαν προ-εκλαμψία και στις περιπτώσεις των γυναικών που εμφάνισαν. Κάτι που δεν αναφέρθηκε πιο πάνω είναι ότι η επιπλοκή της προ-εκλαμψίας πιστεύεται ότι μπορεί να οφείλεται στην μη κανονική ανάπτυξη του πλακούντα. Αυτό περιλαμβάνει την μη φυσιολογική 'εισβολή' τροφοβλαστών (κύτταρα που παρέχουν τροφή στο έμβρυο και εξελίσσονται σε ένα μεγάλο τμήμα του πλακούντα) των σπειροειδή αρτηριών της μήτρας. Περιλαμβάνει επίσης και τη μείωση της αγγειακής αντίστασης του αίματος στην μητροπλακούντια κυκλοφορία. Το υπερηχογράφημα Doppler σε αρτηρίες στη μήτρα μπορεί να αποτιμήσει την μητροπλακούντια κυκλοφορία και από τις κυματομορφές να υπολογιστεί ο δείκτης ταχύτητας της ροής του αίματος (pulsatility index). Το pulsatility index όταν έχει μεγάλες τιμές αυξάνει την πιθανότητα εμφάνισης προ-εκλαμψίας στις εγκύους γυναίκες. [31] Στις πιο πάνω γραφικές ο μέσος όρος στην

τιμή του Doppler test στις περιπτώσεις των γυναικών χωρίς προ-εκλαμψία ήταν 1.6, ενώ στις περιπτώσεις των γυναικών που εμφάνισαν προ-εκλαμψία ήταν σχεδόν 2. Ακόμα παρατηρείται ότι στην δεύτερη γραφική υπάρχουν περιπτώσεις γυναικών με αρκετά μεγάλες τιμές στο Doppler test και εμφάνισαν προ-εκλαμψία, κάτι που δεν παρατηρείται στην πρώτη γραφική.

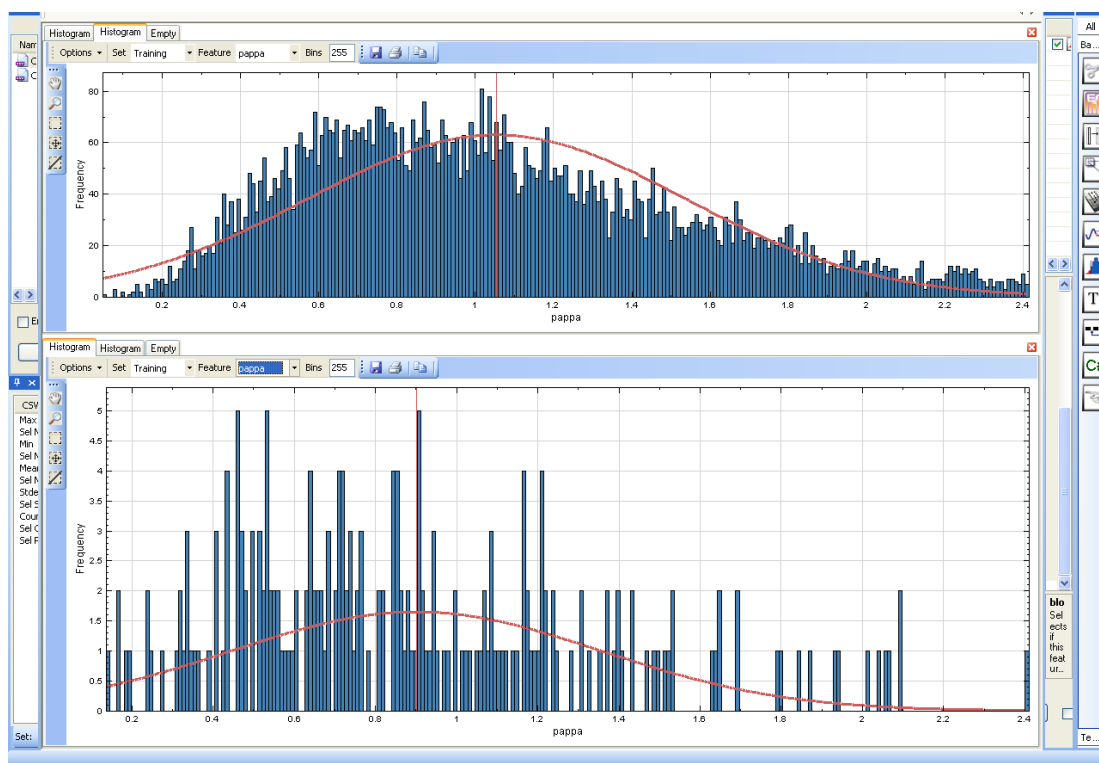
Στην γραφική παράσταση του σχήματος 4.3.10 βλέπουμε τη συσχέτιση του παράγοντα Doppler με τους υπόλοιπους παράγοντες. Παρατηρείται θετική συσχέτιση μεταξύ του παράγοντα αυτού και των output PEarly <34w, output PE middle 34 – 37w, output PE late 38-42w, race Black, previous pregnancies. Αυτό επιβεβαιώνει ότι όσο πιο αυξημένη είναι η τιμή του Doppler test, η έγκυος γυναίκα εμφανίζει προ-εκλαμψία.



Σχήμα 4.3.10 Γραμμική συσχέτιση παράγοντα Doppler με τους υπόλοιπους παράγοντες

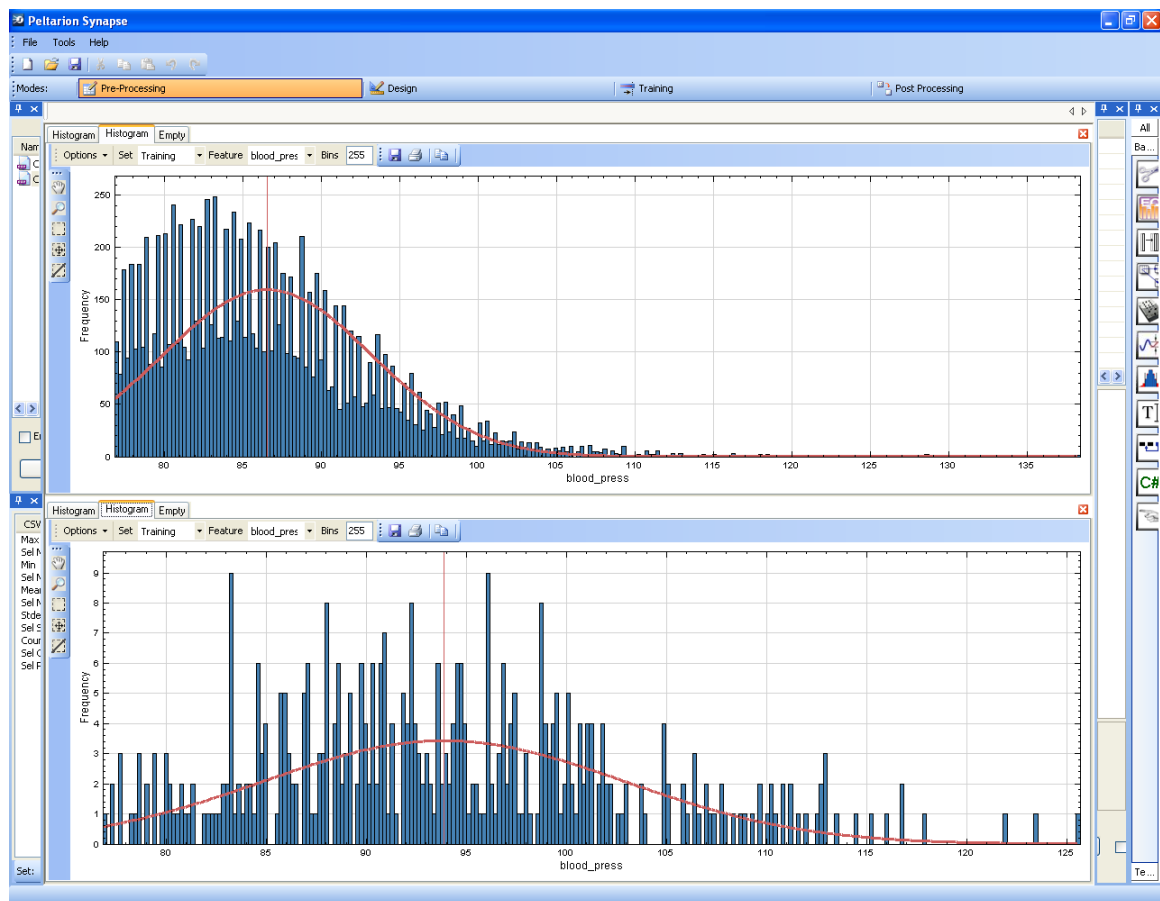


Σχήμα 4.3.11 Ιστόγραμμα τιμών παράγοντα bhCG στις περιπτώσεις που δεν εμφάνισαν προ-εκλαμψία (πρώτο ιστόγραμμα) και στις περιπτώσεις που εμφάνισαν (δεύτερο ιστόγραμμα).



Σχήμα 4.3.12 Ιστόγραμμα τιμών παράγοντα Pappa στις περιπτώσεις που δεν εμφάνισαν προ-εκλαμψία (πρώτο ιστόγραμμα) και στις περιπτώσεις που εμφάνισαν (δεύτερο ιστόγραμμα).

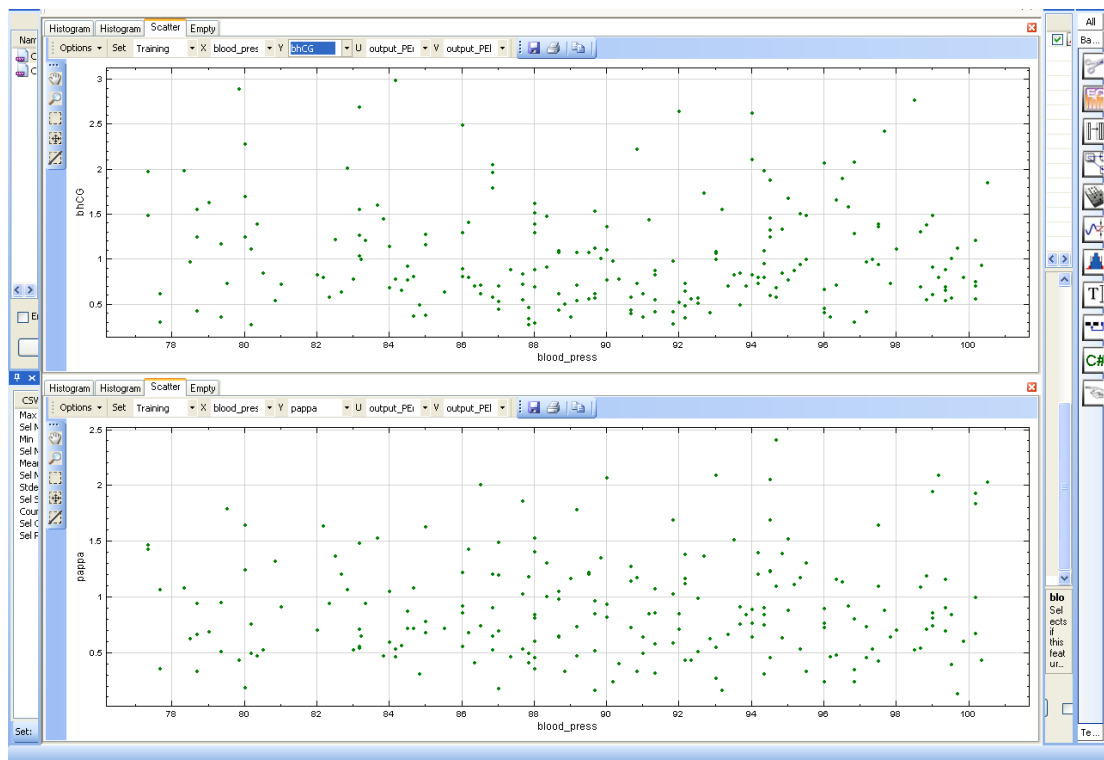
Τα σχήματα 4.3.11 και 4.3.12 είναι ιστογράμματα που παρουσιάζουν τις τιμές της ορμόνης bhCG και της πρωτεΐνης PAPPA στις περιπτώσεις των εγκύων γυναικών που δεν εμφάνισαν ή εμφάνισαν προ-εκλαμψία. Αυτά προέκυψαν και μετά τη χρήση του εργαλείου που αφαιρεί δεδομένα εκτός του κανονικού πεδίου τιμών. Στο ιστόγραμμα των τιμών της ορμόνης bhCG ο μέσος όρος για τις γυναίκες που δεν εμφάνισαν ήταν 1.069 και για τις γυναίκες που εμφάνισαν ήταν 1.021. Στο ιστόγραμμα των τιμών της πρωτεΐνης PAPPA ο μέσος όρος για τις γυναίκες που δεν εμφάνισαν ήταν 1.05 και για τις γυναίκες που εμφάνισαν ήταν 0.90. Ο μέσος όρος και στις δύο περιπτώσεις δεν έχει σημαντική διαφορά. Έτσι δεν μπορούμε να εξάγουμε κάποιο συμπέρασμα για τα επίπεδα των δύο παραγόντων στις περιπτώσεις που εμφάνισαν προ-εκλαμψία.



Σχήμα 4.3.13 Ιστόγραμμα τιμών αρτηριακής πίεσης γυναικών που δεν εμφάνισαν προ-εκλαμψία (πρώτο ιστόγραμμα) και περιπτώσεις γυναικών που εμφάνισαν (δεύτερο ιστόγραμμα).

Τα ιστογράμματα στο σχήμα 4.3.13 παρουσιάζουν τις τιμές της αρτηριακής πίεσης στις περιπτώσεις των γυναικών που δεν εμφάνισαν προ-εκλαμψία και στις περιπτώσεις των γυναικών που εμφάνισαν.

Στις περιπτώσεις που δεν εμφάνισαν προ-εκλαμψία παρατηρήθηκε ότι ο μέσος όρος ήταν 86, ενώ στις περιπτώσεις οι οποίες εμφάνισαν προ-εκλαμψία ο μέσος όρος ήταν 94. Ακόμα παρατηρείται ότι στην δεύτερη γραφική, υπάρχουν περιπτώσεις γυναικών με αρκετά μεγάλη τιμή στην αρτηριακή τους πίεση και εμφάνισαν προ-εκλαμψία. Έτσι επιβεβαιώνεται ότι οι γυναίκες με υψηλή αρτηριακή πίεση εμφανίζουν προ-εκλαμψία.



Σχήμα 4.3.14 Καρτεσιανό διάγραμμα μεταξύ αρτηριακής πίεσης και bhCG (πρώτη γραφική) και αρτηριακής πίεσης με PAPP-A.

Η πρώτη γραφική είναι ένα καρτεσιανό διάγραμμα, όπου στον άξονα X είναι ο παράγοντας της αρτηριακής πίεσης και στον Y άξονα είναι η ορμόνη bhCG στις περιπτώσεις των γυναικών που εμφάνισαν προ-εκλαμψία. Στην δεύτερη γραφική είναι πάλι στον άξονα X είναι ο παράγοντας της αρτηριακής πίεσης και στον Y άξονα είναι η πρωτεΐνη PAPP-A στις περιπτώσεις των γυναικών που εμφάνισαν προ-εκλαμψία. Προσπαθήσαμε να δούμε με την αναπαράσταση των δύο αυτών γραφικών αν υπάρχει κάποια σχέση μεταξύ της αρτηριακής πίεσης και των bhCG, και PAPP-A χωρίς όμως να μπορούμε να εξάγουμε κάποια σχέση αφού τα σημεία μας είναι ομοιόμορφα κατανεμημένα στο χώρο προβολής.

Κεφάλαιο 5

Υλοποίηση – Αποτελέσματα

5.1 Επεξεργασία και Κανονικοποίηση Δεδομένων	86
5.2 Προσέγγιση εκπαίδευσης μοντέλου ΑΓΧ με Γενετικούς Αλγορίθμους	90
5.3 Προσέγγιση εκπαίδευσης μοντέλου ΑΓΧ με τον αλγόριθμο NHL	95
5.4 Εκτέλεση εκπαίδευσης και Αποτελέσματα	98
5.5 Επιπρόσθετες Δοκιμές Εκτέλεσης και Αποτελέσματα	101

Στο κεφάλαιο αυτό παρουσιάζεται η μεθοδολογία που ακολουθήθηκε για την εκπαίδευση του μοντέλου του ΑΓΧ και τα αποτελέσματα αυτής της εκπαίδευσης. Περιγράφω την προσέγγιση εκπαίδευσης του μοντέλου του ΑΓΧ με τη χρήση Γενετικών Αλγορίθμων και την προσέγγιση εκπαίδευσης με τον αλγόριθμο NHL. Τέλος παρουσιάζονται τα αποτελέσματα της εκπαίδευσης αυτής.

5.1 Επεξεργασία και Κανονικοποίηση Δεδομένων

Στο προηγούμενο κεφάλαιο αναφέρθηκαν οι παράγοντες που επηρεάζουν στην εμφάνιση της προ-εκλαμψίας σε μια έγκυο γυναίκα καταλήγοντας τελικά στην επιλογή 14 παραγόντων που θα απαρτίζουν και τους κόμβους του μοντέλου του ΑΓΧ.

Οι 14 αυτοί παράγοντες είναι οι εξής: Output, Mother Age, Race, Height, Conception, Smoking, Previous Pregnancies, Patient's mother had preeclampsia, Chronic hypertension, Blood test Pappa, Blood test β -hCG, Doppler και blood pressure.

Στο αρχείο δεδομένων με τις τιμές των περιπτώσεων των εγκύων γυναικών έγινε κάποια επεξεργασία. Αρχικά οι παράγοντες height και weight συνενώθηκαν σε ένα παράγοντα, του δείκτη μάζας σώματος (body mass index - bmi) και ο οποίος υπολογίζεται με την πιο κάτω εξίσωση:

$$Bmi = \frac{weight (kg)}{(height (m))^2} \quad (41),$$

Έτσι οι τελικοί παράγοντες που αποτελούν τους κόμβους του μοντέλου του ΑΓΧ συνοψίζονται στον Πίνακα 5.1.1.

Παράμετροι μοντέλου ΑΓΧ	
C1	Output
C2	MA
C3	Race
C4	BMI
C5	Conception
C6	Smoking
C7	Previous PET
C8	Patients mother had PE
C9	Chronic hypertension
C10	Blood test PAPPa
C11	Blood test β-hCG
C12	Doppler
C13	Blood pressure

Πίνακας 5.1.1 Παράμετροι (κόμβοι) του μοντέλου του ΑΓΧ

Παρατηρήθηκε στο αρχείο ότι υπήρχαν περιπτώσεις εγκύων γυναικών οι οποίες δεν συμμετείχαν στο blood test Pappa ή δεν είχαν κάνει το doppler test, και σε αυτές τις περιπτώσεις δεν είχαν τιμή σε αυτούς τους δύο παράγοντες. Αυτές οι περιπτώσεις διαγράφηκαν από το αρχείο δεδομένων.

Όπως αναφέραμε στην αρχή στο μοντέλο του ΑΓΧ οι κόμβοι παίρνουν τιμές ενεργοποίησης στο πεδίο [0,1]. Γι αυτό το λόγο το οι τιμές των παραγόντων του αρχείου δεδομένων για όλες τις περιπτώσεις των εγκύων γυναικών έπρεπε να κανονικοποιηθούν στο πεδίο [0,1]. Οι τιμές όλων των περιπτώσεων των εγκύων γυναικών στην ουσία θα αποτελέσουν τα αρχικά επίπεδα ενεργοποίησης τιμών των κόμβων με τις οποίες θα εκτελεστεί ο αλγόριθμος αλληλεπιδράσεων των κόμβων στην προσέγγιση της υλοποίησης μας.

Η κανονικοποίηση η οποία παρουσιάζεται και στον πίνακα 5.1.3 έγινε ως εξής, για την παράμετρο race, δόθηκαν τιμές 1 στις περιπτώσεις εγκύων γυναικών μαύρης φυλής λόγω του ότι γυναίκα μαύρης φυλής έχει μεγαλύτερες πιθανότητες να εμφανίσει προ-εκλαμψία, με το ίδιο σκεπτικό δόθηκαν τιμές 0.4 και 0.2 στις περιπτώσεις εγκύων γυναικών φυλής South Asian και Mixed αντίστοιχα. Αφού όπως φάνηκε από τα δεδομένα στο αρχείο γυναίκες με τις συγκεκριμένες φυλές είχαν θετική συσχέτιση με τον παράγοντα του output και άρα είχαν εμφανίσει προ-εκλαμψία. Για τον παράγοντα previous pregnancies δόθηκαν τιμές 1 όταν η έγκυος γυναίκα είχε εμφανίσει σε προηγούμενες της εγκυμοσύνης προ-εκλαμψία και άρα το ρίσκο να ξαναεμφανίσει είναι μεγάλο, δόθηκε τιμή 0.5 όταν είναι η πρώτη της εγκυμοσύνη γιατί όπως αναφέρει η βιβλιογραφία όταν είναι η πρώτη της εγκυμοσύνη μιας γυναίκας έχει περισσότερες πιθανότητες να εμφανίσει προ-εκλαμψία και τέλος δίνεται τιμή 0 όταν η έγκυος γυναίκα είχε φυσιολογικές προηγούμενες εγκυμοσύνες χωρίς προ-εκλαμψία. Με τον ίδιο τρόπο δόθηκαν τιμές 0 και 1 στους παράγοντες conception, smoking, patients mother had PET, chronic hypertension. Για την παράμετρο conception δόθηκε τιμή 1 στις περιπτώσεις που είχαν τιμή Ovulation with drugs λόγω της θετικής συσχέτισης του παράγοντα αυτού με το output που υπονοεί ότι γυναίκες που έπαιρναν χάπια για να προκαλέσουν ωορρηξία εμφάνισαν προ-εκλαμψία, και 0 στις περιπτώσεις που είχαν τιμή Spontaneous. Επίσης σε όλες τις περιπτώσεις που η έγκυος γυναίκα είχε τιμή Smoker στην παράμετρο smoking δώσαμε τιμή 1. Στην παράμετρο previous Pet, δόθηκε τιμή 1 στις περιπτώσεις που είχαν τιμή Previous pregnancies with PE και 0 στις περιπτώσεις που είχαν τιμή Previous pregnancies without PE και No previous pregnancies. Για τις παραμέτρους patients mother had PET και chronic hypertension δόθηκε τιμή 1 όταν αυτές είχαν τιμές Patients mother had PE και Chronic hypertension αντίστοιχα, ενώ δόθηκε η τιμή 0 όταν αυτές είχαν τιμή No. Για τις υπόλοιπες παραμέτρους, MA, BMI, blood test PAPPa, Blood test β-hCG, Doppler και blood pressure οι τιμές τους υπολογίστηκαν με την εξής φόρμουλα:

$$Normalized_Value = \frac{value}{max_parameter_value} \quad (42),$$

όπου,

value, η τιμή για την συγκεκριμένη παράμετρο

max_parameter_value, η μέγιστη τιμή που έχει κάθε συγκεκριμένη παράμετρος.

Οι μέγιστες τιμές που παίρνουν οι παράμετροι παρουσιάζονται στον πίνακα 5.1.2

Παράμετροι μοντέλου ΑΓΧ	Μέγιστη τιμή παραμέτρου
MA	49.6
BMI	61.27
Blood Test PAPP	7.87
Blood Test β -hCG	19.92
Doppler	4.83
Blood Pressure	138.33

Πίνακας 5.1.2 Πίνακας με τις μέγιστες τιμές παραμέτρων για τις παραμέτρους που έχουν τιμή διαφορετική από 0 και 1.

ACTIVATION LEVEL NORMALIZATION 1		
	Τιμή αρχείου	Κανονικοποιημένη Τιμή
Output	1: Normal	0
	2: Abnormal (with Preeclampsia)	1
MA	Range [13.7 49.6]	value/49.6
Race	1: White	0
	2: South Asian	0.4
	3: Black	1
	4: East Asian	0
	5: Mixed	0.2
BMI (Weight/height ²)	Range [15.2 61.3]	value/61.3
Conception	1: Spontaneous	0
	2: Ovulation Drugs	1
Smoking	0: No	0
	1: Smoker	1
Previous Pregnancies	0: No previous pregnancies	0.5
	1: Previous pregnancies without PET	0
	2: Previous Pregnancies with PET	1

Patients mother had PET	0: No	0
	1: Patients mother had PET	1
Chronic Hypertension	0: No	0
	1: Chronic hypertension	1
Blood test PAPPa	Range [0.047 7.870]	value/7.870
Blood test bhcg	Range [0.04 19.92]	value/19.92
Doppler	Range [0.32 4.83]	value/4.83
BP	Range [49.17 138.33]	value/138.33

Πίνακας 5.1.3 Κανονικοποίηση στο πεδίο [0,1] των τιμών καταστάσεων των κόμβων για όλες τις τιμές των παραμέτρων του αρχείου

Από το αρχικό αρχείο δεδομένων δημιουργήθηκε ένα αρχείο, το οποίο περιλαμβάνει τις 13 παραμέτρους. Από αυτό το αρχείο στην excel επιλέγεται ένα σύνολο 500 περιπτώσεων, από τις οποίες οι 250 είχαν εμφανίσει προ-εκλαμψία (δηλαδή είχαν στο output τιμή 1) και οι υπόλοιπες 250 δεν είχαν εμφανίσει προ-εκλαμψία (δηλαδή είχαν στο output τιμή 0). Οι 500 περιπτώσεις επιλέγηκαν τυχαία, και επιλέχθηκε ίσος αριθμός περιπτώσεων που εμφάνισαν ή όχι προ-εκλαμψία για να εκπαιδευτεί σωστά το μοντέλο του ΑΓΧ.

5.2 Προσέγγιση εκπαίδευσης μοντέλου ΑΓΧ με Γενετικούς Αλγορίθμους

Για την εκπαίδευση του μοντέλου με τη χρήση γενετικών αλγορίθμων έγινε τροποποίηση του κώδικα στη MATLAB του γενικού γενετικού αλγόριθμου.

Ο χρήστης για την εκτέλεση του κώδικα χρειάζεται να δώσει ως παράμετρο τον μέγιστο αριθμό των χρωμοσωμάτων που θέλει να δημιουργούνται σε κάθε γενιά, τον αριθμό των επαναλήψεων, δηλαδή τον μέγιστο αριθμό γενεών που θα δημιουργηθούν, και τον αριθμό των παραγόντων του μοντέλου του ΑΓΧ.

Το αποτέλεσμα της εκπαίδευσης του μοντέλου του ΑΓΧ είναι ο πίνακας των βαρών για το μοντέλο μας και δηλαδή οι σχέσεις μεταξύ των κόμβων. Αυτός ο πίνακας των βαρών αποτελεί την βέλτιστη λύση ανάμεσα σε πολλές υποψήφιες λύσεις (πίνακες βαρών) που δημιουργούνται σε κάθε γενιά.

Στόχος ήταν να βρεθεί ο πίνακας των βαρών με τον οποίο αν γίνει προσομοίωση του μοντέλου του ΑΓΧ με μια δεδομένη αρχική κατάσταση ενεργοποίησης κόμβων θα προβλέπει σωστά αν η έγκυος γυναίκα θα εμφανίσει ή όχι προ-εκλαμψία.

Στο σημείο αυτό σημειώνεται ότι στο μοντέλο του ΑΓΧ θεωρούμε ότι δεν υπάρχει σχέση από μια παράμετρο προς τον εαυτό της και επίσης ότι η παράμετρος output δεν επηρεάζει τις υπόλοιπες παραμέτρους και άρα δεν υπάρχει ακμή από αυτήν προς τους υπόλοιπους κόμβους.

Όπως αναφέραμε και στο Κεφάλαιο 4 χρειάζεται να γίνει κωδικοποίηση του χρωμοσώματος το οποίο περιέχει πληροφορίες για την λύση του προβλήματος και στην περίπτωση μας, αποτελεί τις τιμές του πίνακα των βαρών του μοντέλου του ΑΓΧ.

Οι μεταβλητές των βαρών είναι 169 στο σύνολο τους, αφού οι κόμβοι μας είναι 13 (13X13). Τα βάρη της διαγωνίου που είναι μηδενικά και της πρώτης γραμμής που είναι τα βάρη των ακμών από την παράμετρο του output προς τις υπόλοιπες παραμέτρους και άρα είναι και αυτά μηδενικά αφαιρέθηκαν, και έτσι μένουν για κωδικοποίηση 144 (12X12) μεταβλητά βάρη. Το χρωμόσωμα κωδικοποιείται σε ένα διάνυσμα δυαδικής μορφής και το μήκος του επιλέγεται να είναι 12X12X4=576.

Ο αριθμός 4, είναι το ελάχιστο μήκος γονιδίων που χρειάζεται για να κωδικοποιηθεί μια μεταβλητή βάρους και επιλέγεται έτσι ώστε να ικανοποιείται η πιο κάτω συνθήκη:

$$(b - a) \times 10^q \leq 2^m \quad (43),$$

όπου

a= -1, b=1, τα όρια του πεδίου τιμών των βαρών μας

q=1, η ακρίβεια δεκαδικού που θέλουμε

και άρα m=4

Οι 4 αυτές θέσεις παίρνουν τιμές 0 ή 1. Άρα η κωδικοποίηση του χρωμοσώματος μας επιτυγχάνεται με ένα δυαδικό (0 ή 1) διάνυσμα 576 ψηφίων.

Ο χρήστης δίνει ως παράμετρο τον μέγιστο αριθμό χρωμοσωμάτων ο οποίος συμβολίζει τις υποψήφιες λύσεις οι οποίες θα δημιουργούνται σε κάθε γενιά. Όλες οι υποψήφιες λύσεις που δημιουργούνται για κάθε γενιά αποθηκεύονται σε ένα δυσδιάστατο πίνακα ο οποίος έχει τόσες γραμμές όσες και ο μέγιστος αριθμός χρωμοσωμάτων και τόσες στήλες όσες το μήκος του χρωμοσώματος. Ο πίνακας αυτός αρχικοποιείται με τυχαίο τρόπο.

Όπως αναφέρθηκε στο Κεφάλαιο 4, ένας σημαντικό στοιχείο για την εφαρμογή των γενετικών αλγορίθμων είναι η συνάρτηση καταλληλότητας. Δηλαδή ο τρόπος με τον οποίο αξιολογούνται οι υποψήφιες λύσεις για το μοντέλο του ΑΓΧ για βρεθεί τελικά η καλύτερη λύση. Στην δική μας περίπτωση και με βάση τα δεδομένα που είχαμε, για μια αρχική κατάσταση ενεργοποίησης των παραμέτρων (κόμβων) γνωρίζαμε το τελικό αποτέλεσμα.

Οι αρχικές καταστάσεις ενεργοποίησης ουσιαστικά είναι οι τιμές των 12 παραμέτρων που θεωρούνται ότι επηρεάζουν για την εμφάνιση της προ-εκλαμψίας και η 13^η παράμετρος η έξοδος. Η έξοδος είναι είτε 1 είτε 0, δηλαδή για τις συγκεκριμένες τιμές ενεργοποίησης των υπόλοιπων 12 παραμέτρων η γυναίκα εμφάνισε ή όχι προ-εκλαμψία.

Σε κάθε γενιά (επανάληψη) του γενετικού αλγορίθμου εκτελείται ο αλγόριθμος προσομοίωσης αλληλεπιδράσεων του ΑΓΧ για όλες τις καταστάσεις ενεργοποίησης που υπάρχουν στο αρχείο δεδομένων που δημιουργήθηκε και για κάθε ένα από τα υποψήφια χρωμοσώματα (πίνακα βαρών) κάθε γενιάς.

Στο τέλος της εκτέλεσης του αλγορίθμου προσομοίωσης αλληλεπιδράσεων του ΑΓΧ συγκρίνουμε την τιμή ενεργοποίησης που έχει η παράμετρος (κόμβος) output με την τιμή που θα έπρεπε να είχε, αφού αυτή όπως αναφέραμε είναι δεδομένη για κάθε περίπτωση. Η τιμή καταλληλότητας που δίνεται σε κάθε χρωμόσωμα είναι το

άθροισμα των λαθών περιπτώσεων που προέβλεψε. Το καλύτερο χρωμόσωμα που θα προκύψει θα είναι αυτό με το μικρότερη τιμή λάθους. Έτσι σε κάθε χρωμόσωμα κάθε γενιάς ανατίθεται μια τιμή καταλληλότητας ανάλογα με το πόσες περιπτώσεις βρήκε λάθος από το σύνολο των δεδομένων.

Σημειώνεται ότι στο σημείο που γίνεται η εκτέλεση του αλγορίθμου προσομοίωσης των αλληλεπιδράσεων του ΑΓΧ, γίνεται χρήση της νέα μέθοδο που προτάθηκε στο άρθρο [2] και αναφέρθηκε στο υποκεφάλαιο 2.5. Στην προσέγγιση αυτή για τον υπολογισμό των νέων τιμών των κόμβων χρειάζεται να υπολογιστεί η διαφορά της τιμής ενεργοποίησης του κόμβου στο προηγούμενο βήμα με την τιμή ενεργοποίησης στο παρών βήμα που στη συνέχεια πολλαπλασιάζεται με το βάρος. Έτσι για το πρώτο επαναληπτικό βήμα του αλγορίθμου, χρειάζεται να οριστεί μια αρχική κατάσταση τιμών των κόμβων, η οποία θεωρούμε ότι αντικατοπτρίζει τις τιμές που έχει μια έγκυος γυναίκα σε φυσιολογική εγκυμοσύνη.

Στον πίνακα 5.2.1 παρουσιάζονται οι τιμές που καθορίστηκαν ως φυσιολογικές τιμές σε μια εγκυμοσύνη για τους παράγοντες του μοντέλου μας. Για τον καθορισμό των φυσιολογικών τιμών των παραμέτρων Blood test PAPP, Blood test β -hCG, Doppler υπολογίστηκε ο μέσος όρος των περιπτώσεων των γυναικών που δεν εμφάνισαν προεκλαμψία από το αρχείο δεδομένων.

Παράμετρος μοντέλου ΑΓΧ	Φυσιολογική τιμή	Κανονικοποιημένη τιμή
Output	Normal	0
MA	20 - 35	0.55
Race	White, South Asian, East Asian, Mixed	0
BMI	18.5 - 25	0.4
Conception	Spontaneous	0
Smoking	No	0
Previous PET	No	0
Patients mother had PE	No	0

Chronic hypertension	No	0
Blood test PAPP-A	1,69	0.2
Blood test β -hCG	1.20	0.6
Doppler	1.69	0.34
Blood pressure	70 - 110	0.65

Πίνακας 5.2.1 Τιμές παραμέτρων που έχει μια έγκυος γυναίκα σε φυσιολογική εγκυμοσύνη

Στη συνέχεια της εκτέλεσης του γενετικού αλγορίθμου ακολουθεί η εξέλιξη του πληθυσμού με την εφαρμογή μηχανισμών, του ανά-συνδυασμού γονιδίων (crossover), της μετάλλαξης (mutation), της επιλογής για την παραγωγή νέου πληθυσμού που θα αποτελεί την επόμενη γενιά. Η εξέλιξη του πληθυσμού γίνεται μέχρι να παραχθεί ο μέγιστος αριθμός γενεών που έθεσε στην αρχή ο χρήστης.

Για τον ανά-συνδυασμό των γονιδίων επιλέγονται με τον γενετικό τελεστή Roulette – Wheel Selection, τα χρωμοσώματα βάση της τιμής καταλληλότητας τους που θα αποτελέσουν τους γονείς για την δημιουργία των νέων απογόνων. Η μέθοδος αυτή υπολογίζει μια πιθανότητα για κάθε ένα από τα χρωμοσώματα στον πληθυσμό βάση της τιμής καταλληλότητας που έχει το κάθε χρωμόσωμα. Η πιθανότητα υπολογίζεται ως ο λόγος της τιμής καταλληλότητας ως προς το άθροισμα όλων των τιμών καταλληλότητας των χρωμοσωμάτων στον πληθυσμό. Η επιλογή γίνεται με τυχαίο τρόπο όπως γυρίζει ο τροχός της ρουλέτας. Τα χρωμοσώματα με την καλύτερη τιμή καταλληλότητας έχουν περισσότερες πιθανότητες να επιλεγούν.

Στην συνέχεια αφού επιλέγηκαν οι γονείς, γίνεται ανά-συνδυασμός (crossover) των χρωμοσωμάτων. Ο ανά-συνδυασμός γίνεται με τον γενετικό τελεστή ενός σημείου ανά-συνδυασμός (one point crossover). Δίνεται μια πιθανότητα με τυχαίο τρόπο σε όλα τα χρωμοσώματα (γονείς) και αν αυτή είναι μικρότερη του $P_c = 0.7$ (70%) που καθορίστηκε ως η πιθανότητα ανά-συνδυασμού τότε πραγματοποιείται ο ανά-συνδυασμός αλλιώς τα χρωμοσώματα περνούν στον επόμενο πληθυσμό ακριβώς όπως είναι. Ο τρόπος με τον οποίο γίνεται ο ανά-συνδυασμός είναι με την επιλογή μιας τυχαίας θέσης που μοιράζει έτσι τα δύο χρωμοσώματα γονείς σε δύο μέρη. Ο πρώτος απόγονος σχηματίζεται από το πρώτο μέρος του πρώτου γονιού και το δεύτερο μέρος

του δεύτερου γονιού. Ενώ ο δεύτερος σχηματίζεται από το πρώτο μέρος του δεύτερου γονιού και το πρώτο μέρος του δεύτερου γονιού.

Ακολούθως εφαρμόζεται ο τελεστής της μετάλλαξης (mutation). Η μετάλλαξη πραγματοποιείται στα χρωμοσώματα με πιθανότητα $P_m = 0.01$ (1%), και αλλάζουν σε τυχαίες θέσεις τα ψηφία από 0-1 και αντίστροφα.

Τέλος χρησιμοποιείται ξανά ο γενετικός τελεστής Best Individuals για την δημιουργία του νέου πληθυσμού. Έτσι επιτυγχάνεται η συμμετοχή των καλύτερων χρωμοσωμάτων (γονέων) στην επόμενη γενεά.

5.3 Προσέγγιση εκπαίδευσης μοντέλου ΑΓΧ με τον αλγόριθμο NHL

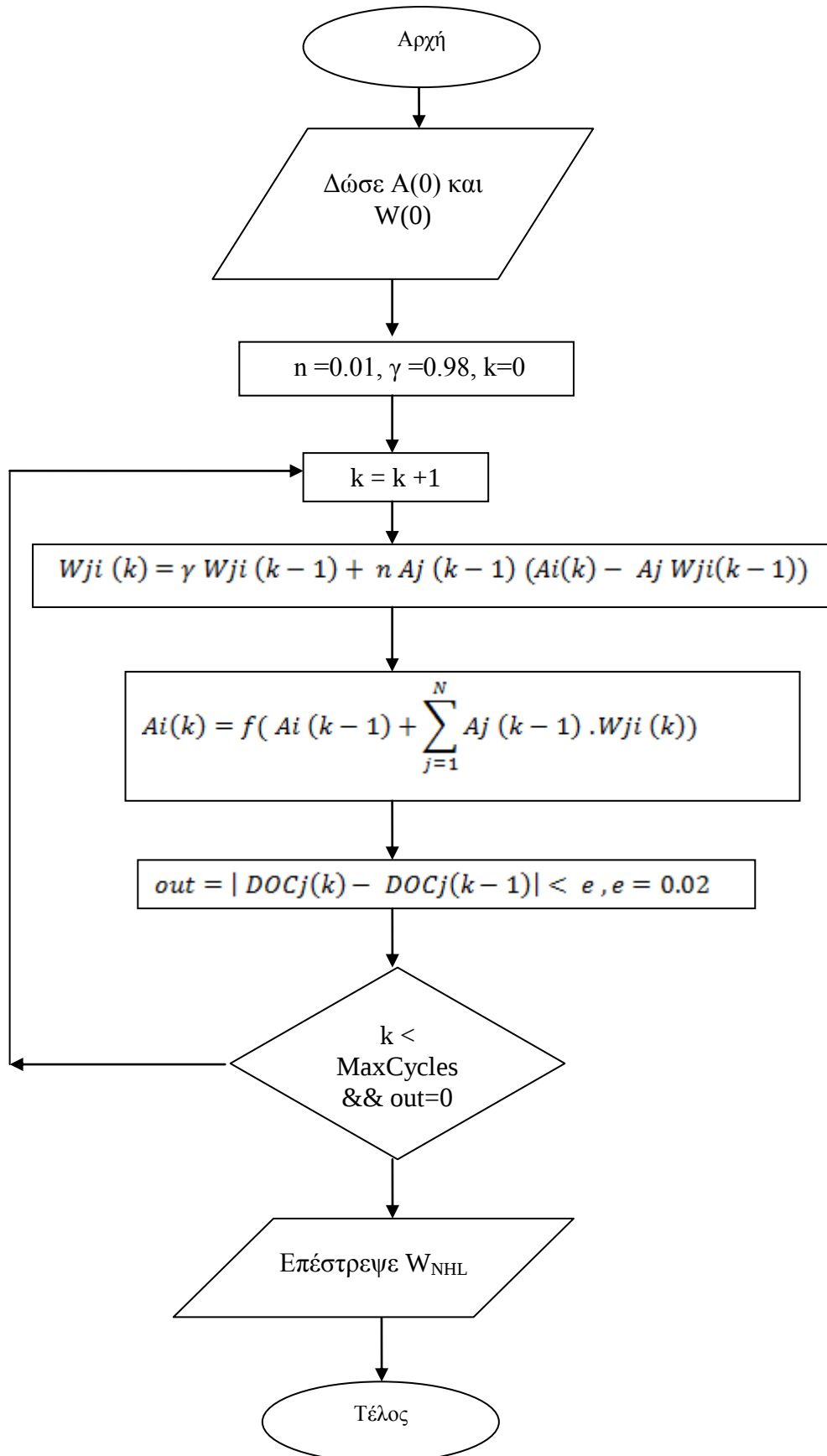
Ο αλγόριθμος NHL που περιγράψαμε στο υποκεφάλαιο 3.2 υλοποιήθηκε στην γλώσσα προγραμματισμού MATLAB. Ο λόγος που επιλέχθηκε η γλώσσα προγραμματισμού MATLAB είναι λόγω του ότι μπορεί εύκολα να δημιουργήσει και να επεξεργαστεί μαθηματικούς πίνακες.

Συγκεκριμένα έχει υλοποιηθεί μια συνάρτηση η οποία δέχεται τρεις παραμέτρους, ένα μονοδιάστατο πίνακα για την αρχική κατάσταση των κόμβων, ένα δυσδιάστατο τετραγωνικό πίνακα για τα αρχικά βάρη τα οποία θέτουν για το μοντέλο του ΑΓΧ οι έμπειροι και ένα μονοδιάστατο πίνακα για τους κόμβους εξόδου επιθυμητής τιμής.

Στην συγκεκριμένη περίπτωση ο αρχικός πίνακας βαρών που θα δινόταν ως είσοδος θα ήταν το αποτέλεσμα της καλύτερης λύσης που προέκυψε από την εκπαίδευση του μοντέλου του ΑΓΧ με τους γενετικούς αλγόριθμους που περιγράψαμε στο υποκεφάλαιο 5.2 πιο πάνω. Αυτό γιατί το μοντέλο δεν καθορίστηκε από έμπειρους.

Οι τιμές του πίνακα των βαρών αλλάζουν μέσω ενός διπλού βρόγχου, ο οποίος περνά απ όλες τις θέσεις του δυσδιάστατου πίνακα και αλλάζει τις τιμές σύμφωνα με την εξίσωση (24) στο κεφάλαιο 3. Αρχικά ο αλγόριθμος υλοποιήθηκε να αλλάζει όλες τις τιμές των βαρών εκτός της διαγωνίου και επαληθεύτηκε με το παράδειγμα του άρθρου [13]. Στη συνέχεια πρόσθεσα ακόμα μια συνθήκη η οποία αφήνει μηδέν τις τιμές της

πρώτης γραμμής του πίνακα, αφού ο κόμβος output δεν έχει ακμές προς τους υπόλοιπους κόμβους του μοντέλου. Στην εξίσωση της αλλαγής των βαρών όπως αναφέραμε και στην περιγραφή του αλγορίθμου υπάρχουν δύο παράμετροι, ο ρυθμός εκμάθησης η και ο συντελεστής μείωσης βάρους γ . Οι τιμές των παραμέτρων αυτών καθορίζονται στην αρχή της συνάρτησης. Οι τιμές οι οποίες επιλέγηκαν είναι $\eta=0.01$ και $\gamma=0.98$. Αφού αλλάξουν οι τιμές των βαρών υπολογίζονται οι τιμές καταστάσεων των κόμβων σύμφωνα με την εξίσωση (22) του κεφαλαίου 3. Η τιμή αυτή περνά από την σιγμοειδή συνάρτηση έτσι ώστε να κυμαίνεται στο πεδίο τιμών $[0,1]$. Η σιγμοειδής συνάρτηση υλοποιήθηκε ως ξεχωριστή συνάρτηση η οποία δέχεται μια τιμή και επιστρέφει το αποτέλεσμα. Για τον τερματισμό του αλγορίθμου ελέγχεται αν ικανοποιείται η συνθήκη (29) που αναφέρεται στο Κεφάλαιο 3 και εάν εκτελεστεί ένας μέγιστος αριθμός επαναλήψεων ο οποίος ορίζεται ως 1000. Η συνθήκη τερματισμού υλοποιείται ως ξεχωριστή συνάρτηση η οποία δέχεται τρεις παραμέτρους, τον πίνακα με τις παλιές τιμές καταστάσεων των κόμβων, τον πίνακα με τις νέες τιμές των κόμβων που υπολογίστηκαν μετά την αλλαγή των τιμών των βαρών και τον πίνακα με τους κόμβους εξόδου επιθυμητής τιμής. Στην συνάρτηση αυτή γίνεται έλεγχος για όλους τους κόμβους εξόδους επιθυμητής τιμής την διαφορά της νέας τιμής με την παλιά αν είναι μικρότερη από μια τιμή ϵ . Η τιμή αυτή καθορίστηκε να είναι 0.02. Τέλος αν δεν ικανοποιούνται οι δύο συνθήκες συνεχίζεται η εκτέλεση του αλγορίθμου με τις νέες τιμές του πίνακα βαρών και τις νέες τιμές του πίνακα καταστάσεων των κόμβων μέχρι να ικανοποιηθούν οι συνθήκες. Όταν οι συνθήκες ικανοποιηθούν επιστρέφεται ο πίνακας με τις τιμές των βαρών.



Σχήμα 5.3.1 Λογικό διάγραμμα της υλοποίησης του αλγορίθμου NHL

5.4 Εκτέλεση εκπαίδευσης και αποτελέσματα

Στόχος της εκπαίδευσης του μοντέλου του ΑΓΧ ήταν να βρεθεί ο πίνακας των βαρών, δηλαδή οι σχέσεις αιτιότητας και ο βαθμός αιτιότητας μεταξύ των κόμβων. Με αυτό τον τρόπο θα προκύψει το μοντέλο του ΑΓΧ το οποίο δίνοντας μια οποιαδήποτε αρχική κατάσταση ενεργοποίησης στους κόμβους που αντιστοιχεί στις μετρήσεις από την εξέταση μιας εγκύου γυναίκας θα προβλέπει εάν αυτή η γυναίκα θα εμφανίσει ή όχι προ-εκλαμψία.

Αρχικά έγινε εκπαίδευση του μοντέλου με τους γενετικούς αλγορίθμους όπως περιγράψαμε στο υποκεφάλαιο 5.2. Ακολούθησα την μέθοδο δοκιμής – λάθους εκτελώντας τον αλγόριθμο με διαφορετικό αριθμό μέγιστων χρωμοσωμάτων και επαναλήψεων κάθε φορά. Προσπαθώντας έτσι να πετύχω έναν συνδυασμό που θα μου δίνει την καλύτερη λύση. Οι δοκιμές που έγιναν φαίνονται στον πίνακα 5.4.1.

Μέγιστος Αριθμός Χρωμοσωμάτων	Μέγιστος Αριθμός Επαναλήψεων	Τιμή καταλληλότητας καλύτερης λύσης
1000	200	205
1000	100	207
600	400	204
500	100	207
500	500	204
300	900	204
200	800	204
200	400	205
100	500	205

Πίνακας 5.4.1 Αποτελέσματα εκπαίδευσης με τη χρήση γενετικού αλγορίθμου

Τα αποτελέσματα όπως φαίνεται στον πίνακα δυστυχώς δεν είναι πολύ ενθαρρυντικά. Όπως αναφέραμε στο υποκεφάλαιο 5.2 η τιμή αυτή καθορίζει τις περιπτώσεις λάθους του κάθε χρωμοσώματος. Με την μέθοδο που εφαρμόσαμε με τους γενετικούς αλγορίθμους προσπαθήσαμε να βελτιώσουμε την αρχική λύση και να μειωθεί όσο

περισσότερο το σφάλμα και οι περιπτώσεις λάθους. Δυστυχώς όπως φαίνεται η λύση με την καλύτερη τιμή καταλληλότητας είναι 204. Αυτό πάει να πει ότι ο πίνακας των βαρών προβλέπει σωστά 296 περιπτώσεις από τις 500, και 204 τις προβλέπει λάθος.

Παρατηρείται επίσης ότι με τον ίδιο αριθμό χρωμοσωμάτων και μεγαλύτερο αριθμό επαναλήψεων, δηλαδή στις περιπτώσεις των 1000, 500 και 200 χρωμοσωμάτων, η τιμή καταλληλότητας της καλύτερης λύσης είναι μικρότερη. Επίσης παρατηρήθηκε ότι σχεδόν για όλες τις περιπτώσεις που εκτελέστηκε ο αλγόριθμος, μετά το τέλος της πρώτης επανάληψης η τιμή καταλληλότητας της καλύτερης λύσης ήταν περίπου 250, δηλαδή προέβλεπε λάθος τα μισά δεδομένα αλλά μετά από την ολοκλήρωση των επαναλήψεων η λύση δεν βελτιωνόταν δραματικά.

Στην συνέχεια εκτελέστηκε ο αλγόριθμος NHL, δίνοντας ως είσοδο τον πίνακα των βαρών που προέκυψε από την εκπαίδευση των Γενετικών αλγορίθμων στις περιπτώσεις που η τιμή καταλληλότητας ήταν η καλύτερη. Ως αρχική κατάσταση τιμών ενεργοποίησης κόμβων, δόθηκε ο πίνακας με τις τιμές που έχει μια γυναίκα σε φυσιολογική εγκυμοσύνη και αναφέραμε στο υποκεφάλαιο 2.1 πιο πάνω. Ο κόμβος του οποίου μας ενδιαφέρει η τιμή του είναι ο κόμβος της παραμέτρου output.

Αφού έτρεξε ο αλγόριθμος και μας επέστρεψε τον τροποποιημένο πίνακα βαρών για όλες τις περιπτώσεις, έγινε εκτέλεση του αλγορίθμου προσομοίωσης των αλληλεπιδράσεων του μοντέλου του ΑΓΧ. Ο αλγόριθμος αλληλεπιδράσεων εκτελέστηκε με όλες τις 500 αρχικές καταστάσεις του αρχείου που δημιουργήσαμε και κάθε φορά υπολογίζονταν οι λάθος περιπτώσεις. Τα αποτελέσματα παρουσιάζονται στον πίνακα 5.4.2

	Μέγιστος Αριθμός Χρωμοσωμάτων	Μέγιστος Αριθμός Επαναλήψεων	Τιμή καταλληλότητας λύσης (Λάθος περιπτώσεις)
ΓΑ	600	400	204
NHL			219

ΓΑ	500	500	204
NHL			222

ΓΑ	300	900	204
NHL			216

ΓΑ	200	800	204
NHL			227

Πίνακας 5.4.2 Αποτελέσματα εκπαίδευσης με τον πίνακα που τροποποίησε ο NHL αλγόριθμος για τις καλύτερες λύσεις που έδωσε ο γενετικός αλγόριθμος

Τα αποτελέσματα με τις τροποποιήσεις που έκανε στις τιμές των βαρών ο αλγόριθμος NHL δεν είναι καλύτερα, αφού όπως φαίνεται στον πίνακα μετά την προσομοίωση της αλληλεπίδρασης του μοντέλου οι λάθος περιπτώσεις που προβλέπει το μοντέλο του ΑΓΧ είναι πολύ περισσότερες.

Επίσης οι περιπτώσεις που δεν προβλέπει σωστά το δίκτυο είναι οι περιπτώσεις των γυναικών με προ-εκλαμψία. Για παράδειγμα στην πρώτη περίπτωση του Πίνακα 5.4.2 οι λάθος περιπτώσεις ήταν συνολικά 219, από τις οποίες 217 ήταν οι λάθος περιπτώσεις που εμφάνισαν προ-εκλαμψία και 2 ήταν οι λάθος περιπτώσεις που δεν εμφάνισα προ-εκλαμψία. Το μοντέλο προβλέπει ορθά μόνο 1,2 περιπτώσεις με προ-εκλαμψία σε κάποιες από περιπτώσεις ενώ μπορεί να μην προβλέψει καμιά

5.6 Επιπρόσθετες Δοκιμές Εκτέλεσης και Αποτελέσματα

Στην προηγούμενο υποκεφάλαιο στην προσπάθεια εκμάθησης του μοντέλου του ΑΓΧ με την προσέγγιση που έγινε με τους γενετικούς αλγόριθμους έγιναν διάφορες δοκιμές με διάφορους συνδυασμούς μέγιστου αριθμού χρωμοσωμάτων και μέγιστου αριθμού γενεών. Όμως τα αποτελέσματα που προέκυψαν δεν ήταν καλά λόγω του μεγάλου αριθμού των περιπτώσεων που προέβλεπε λάθος ο αλγόριθμος. Σε αυτό το σημείο έγιναν περαιτέρω προσπάθειες και δοκιμές για να προκύψουν κάποια συμπεράσματα για την συμπεριφορά του γενετικού αλγόριθμου με την αλλαγή διαφόρων παραμέτρων στην εκμάθηση του μοντέλου μας αλλά και για να προκύψουν τυχόν καλύτερα αποτελέσματα.

Αρχικά έγινε μια μικρή διαφοροποίηση στην κωδικοποίηση που αναφέραμε στο υποκεφάλαιο 5.1. Συγκεκριμένα η κανονικοποίηση του παράγοντα της ηλικίας της μητέρας και του δείκτη μάζας σώματος γίνεται διαφορετικά, ενώ οι υπόλοιποι παράγοντες κανονικοποιούνται όπως αναφέρεται στο υποκεφάλαιο 5.1. Για την ηλικία της μητέρας θέτουμε τις εξής κατηγορίες. Αν η ηλικία της είναι μικρότερη των 18 ή μεγαλύτερη και ίση του 37 παίρνει τιμή 1 και αυτό γιατί οι γυναίκες που μένουν έγκυες νωρίς κατά την εφηβεία τους έχουν μεγαλύτερο κίνδυνο να εμφανίσουν προ-εκλαμψία. Αν η ηλικία της μητέρας είναι μεταξύ 25 και 37 παίρνει τιμή 0.3 και δίνεται τιμή 0 στις ηλικίες από 18 μέχρι 25 που θεωρούνται οι πιο καλές ηλικίες για να μείνει μια γυναίκα έγκυος. Για τον παράγοντα του δείκτη μάζας σώματος θέτουμε τις εξής κατηγορίες: Αν η τιμή του παράγοντα αυτού είναι μικρότερη από 25 δίνεται η τιμή 0 λόγω του η συγκεκριμένη περίπτωση εγκύου γυναίκας έχει βάρος χαμηλότερο από το φυσιολογικό ή φυσιολογικό, η περίπτωση να είναι ο δείκτης μάζας σώματος μεταξύ 25 και 30 θεωρείτε ότι η έγκυος γυναίκα είναι υπέρβαρη και δίνεται τιμή 0.5, ενώ τιμές του δείκτη μάζας σώματος μεγαλύτερες του 30 θεωρείτε ότι η έγκυος γυναίκα ανήκει στο 1^ο με 3^ο βαθμού παχυσαρκίας και δίνεται η τιμή 1. Αυτό γιατί από την βιβλιογραφία αναφέρεται ότι οι γυναίκες με ψηλό δείκτη μάζας σώματος έχουν περισσότερες πιθανότητες να εμφανίσουν προ-εκλαμψία. Η νέα κανονικοποίηση παρουσιάζεται στον πίνακα 5.6.1.

ACTIVATION LEVEL NORMALIZATION 2		
	Τιμή αρχείου	Κανονικοποιημένη Τιμή
Output	1: Normal	0
	2: Abnormal (with Preeclampsia)	1
MA	Range [13.7 49.6]	$Ma < 18 \rightarrow 1$ $Ma \geq 37 \rightarrow 1$ $Ma \geq 25$ and $Ma < 37 \rightarrow 0.3$ Else 0
Race	1: White	0
	2: South Asian	0.4
	3: Black	1
	4: East Asian	0
	5: Mixed	0.2
BMI (Weight/height ²)	Range [15.2 61.3]	$BMI < 25 \rightarrow 0$ $BMI \geq 25$ and $BMI \leq 30 \rightarrow 0.5$ Else 1
Conception	1: Spontaneous	0
	2: Ovulation Drugs	1
Smoking	0: No	0
	1: Smoker	1
Previous Pregnancies	0: No previous pregnancies	0.5
	1: Previous pregnancies without PET	0
	2: Previous Pregnancies with PET	1
Patients mother had PET	0: No	0
	1: Patients mother had PET	1
Chronic Hypertension	0: No	0
	1: Chronic hypertension	1
Blood test PAPPa	Range [0.047 7.870]	value/7.870
Blood test bhcg	Range [0.04 19.92]	value/19.92
Doppler	Range [0.32 4.83]	value/4.83
BP	Range [49.17 138.33]	value/138.33

Πίνακας 5.6.1 Νέα Κανονικοποίηση στο πεδίο [0,1] των τιμών καταστάσεων των κόμβων για όλες τις τιμές των παραμέτρων του αρχείου

Στη συνέχεια έγινε η κανονικοποίηση πέντε διαφορετικών αρχείων που αποτελούνταν από διαφορετικές περιπτώσεις εγκύων γυναικών και οι οποίες επιλέχθηκαν από το αρχικό αρχείο. Κάθε αρχείο αποτελείται από 840 περιπτώσεις εγκύων γυναικών από τις οποίες οι 420 δεν εμφάνισαν προ-εκλαμψία και οι υπόλοιπες 420 εμφάνισαν προ-εκλαμψία. Οι περιπτώσεις των εγκύων γυναικών είναι διαφορετικές σε κάθε αρχείο και η επιλογή τους έγινε με τρόπο ώστε να μην υπάρχουν πολλές περιπτώσεις εγκύων γυναικών που να έχουν ίδιες τιμές για κάθε ένα παράγοντα αλλά να υπάρχουν περιπτώσεις εγκύων γυναικών με τιμές από όλο το πεδίο τιμών για κάθε ένα παράγοντα. Οι τιμές των αρχείων κανονικοποιήθηκαν και με τους δύο τρόπους που αναφέρθηκαν πιο πάνω.

Ακολούθως έγιναν διάφορες εκτελέσεις του αλγορίθμου εκμάθησης του μοντέλου με τα αρχεία αυτά. Στους πίνακες που ακολουθούν παρουσιάζω τις διάφορες εκτελέσεις και με τις διάφορες παραμέτρους που εκτελέστηκαν.

Testing	DataSet1 with normalization 1			
Case	1	2	3	4
Max Generations (iterations)	100	100	100	100
Max Chromosomes	500	500	500	500
Crossover Selection Method	Roulette Wheel	Roulette Wheel	Roulette Wheel	Roulette Wheel
Crossover Method	One Point	One Point	One Point	Two Point
Crossover Probability	0.7	0.1	0.1	0.1
Mutation Method	Random	Random	Random	Random
Mutation Probability	0.01	0.01	0.001	0.01
Replacement Method	Best individuals	Best individuals	Best individuals	Best individuals
Best Min Fit (out of 840)	346	344	352	346
Wrong Normal (out of 420)	6	8	9	6
Wrong Pet (out of 420)	340	336	343	340

Πίνακας 5.6.2 Εκτελέσεις Εκμάθησης με σύνολο δεδομένων 1

Παρατηρώντας τον πίνακα 5.6.2, συγκρίνοντας την περίπτωση 1 με την 2 βλέπουμε ότι μειώνοντας την πιθανότητα του crossover από 0.7 σε 0.1, οι περιπτώσεις λάθους είναι

με πάρα πολύ μικρή διαφορά ελάχιστες. Την ίδια διαφορά μπορούμε να δούμε αν συγκρίνουμε τις περιπτώσεις 2 και 4 με διαφορά στη μέθοδο crossover από One σε Two Point. Συγκρίνοντας την περίπτωση 2 με 3 βλέπουμε ότι με τη μείωση της πιθανότητας της μετάλλαξης οι λάθος περιπτώσεις είναι πολύ περισσότερες.

Testing	DataSet1 with normalization 1		DataSet1 with normalization 2	
Case	1	2	3	4
Max Generations (iterations)	100	100	100	375
Max Chromosomes	500	500	500	1000
Crossover Selection Method	Tournament (size=100)	Roulette Wheel	Tournament (size=100)	Tournament (size=100)
Crossover Method	Two Point	Two Point	Two Point	Two Point
Crossover Probability	0.1	0.1	0.1	0.1
Mutation Method	Random	Random	Random	Random
Mutation Probability	0.001	0.01	0.01	0.01
Replacement Method	Best Individuals	Best individuals	Best individuals	Best individuals
Best Min Fit (out of 840)	347	346	299	276
Wrong Normal (out of 420)	6	6	10	
Wrong Pet (out of 420)	341	340	289	

Πίνακας 5.6.3 Εκτελέσεις Εκμάθησης με σύνολο δεδομένων 1 με τις δύο κανονικοποιήσεις

Συγκρίνοντας την περίπτωση 3 και 4 του πίνακα 5.6.3 βλέπουμε ότι διπλασιάσαμε τον μέγιστο αριθμό χρωμοσωμάτων και ότι στις περισσότερες εκτελέσεις που έγιναν οι λανθασμένες περιπτώσεις μειώνονται.

Testing	DataSet2 with normalization 2		
Case	1	2	3
Max Generations (iterations)	100	370	100
Max Chromosomes	500	1000	500
Crossover Selection Method	Roulette Wheel	Roulette Wheel	Roulette Wheel
Crossover Method	One Point	One Point	One Point
Crossover Probability	0.7	0.7	0.9
Mutation Method	Random	Random	Random
Mutation Probability	0.01	0.01	0.009
Replacement Method	Best individuals	Best individuals	Best individuals
Best Min Fit (out of 840)	273	256	281
Wrong Normal (out of 420)	16		15
Wrong Pet (out of 420)	257		266

Πίνακας 5.6.4 Εκτελέσεις Εκμάθησης με σύνολο δεδομένων 2 με την 2^η κανονικοποίηση

Παρατηρείται ότι αν συγκρίνουμε την περίπτωση 1 και 2 του πίνακα 5.6.4 έχοντας διπλάσιο μέγιστο αριθμό χρωμοσωμάτων σε περισσότερες εκτελέσεις που έγιναν οι λανθασμένες περιπτώσεις μειώνονται. Ακόμα και από τις εκτελέσεις του συνόλου δεδομένων 1 και 2 παρατηρείται ότι 2^η κανονικοποίηση δίνει λιγότερες λανθασμένες περιπτώσεις.

Testing	DataSet3 with normalization 1				
Case	1	2	3	4	5
Max Generations (iterations)	100	100	100	100	100
Max Chromosomes	500	500	500	500	500
Crossover Selection Method	Roulette Wheel	Roulette Wheel	Tournament (size=100)	Tournament (size=100)	Tournament (size=100)
Crossover Method	One Point	One Point	Two Point	Two Point	Two Point
Crossover Probability	0.7	0.9	0.7	0.1	0.1
Mutation Method	Random	Random	Random	Random	Random
Mutation Probability	0.01	0.01	0.01	0.01	0.001
Replacement Method	Best individuals	Best individuals	Best individuals	Best individuals	Best individuals
Best Min Fit (out of 840)	343	342	340	344	344
Wrong Normal (out of 420)	4	8	6	5	10
Wrong Pet (out of 420)	339	334	334	339	334

Πίνακας 5.6.5 Εκτελέσεις Εκμάθησης με σύνολο δεδομένων 3 με την 1^η κανονικοποίηση

Συγκρίνοντας τις περιπτώσεις 1 και 2 του σχήματος 5.6.5 παρατηρούμε ότι με πιθανότητα crossover μεγαλύτερη οι λάθος περιπτώσεις είναι ελάχιστα λιγότερες. Το ίδιο παρατηρείται και στις περιπτώσεις 3 και 4 παρόλο που η μέθοδος επιλογής για crossover είναι tournament και επιλέγεται μέθοδος Two point για crossover. Όταν η πιθανότητα crossover είναι μεγαλύτερη οι λάθος περιπτώσεις είναι ελάχιστα λιγότερες. Συγκρίνοντας τις περιπτώσεις 4 και 5 παρατηρούμε ότι αν και οι συνολικές περιπτώσεις λάθους είναι οι ίδιες με μικρότερη πιθανότητα μετάλλαξης οι περιπτώσεις λάθους με PET είναι ελάχιστα λιγότερες.

Testing	DataSet3 with normalization 2		
Case	1	2	3
Max Generations (iterations)	100	100	500
Max Chromosomes	500	500	1000
Crossover Selection Method	Tournament (size=100)	Tournament (size=100)	Tournament (size=100)
Crossover Method	One Point	Two Point	One Point
Crossover Probability	0.1	0.1	0.1
Mutation Method	Random	Random	Random
Mutation Probability	0.01	0.01	0.01
Replacement Method	Best individuals	Best individuals	Best individuals
Best Min Fit (out of 840)	305	307	273
Wrong Normal (out of 420)	13	16	14
Wrong Pet (out of 420)	292	291	259

Πίνακας 5.6.6 Εκτελέσεις Εκμάθησης με σύνολο δεδομένων 3 με την 1^η κανονικοποίηση

Συγκρίνοντας τις περιπτώσεις 1 και 2 του πίνακα 5.6.6 φαίνεται ότι με τη μέθοδο Tournament εφαρμόσουμε One Point crossover το συνολικό λάθος περιπτώσεων είναι ελάχιστα λιγότερο. Συγκρίνοντας τις περιπτώσεις 1 και 3 παρατηρούμε όταν ο μέγιστος αριθμός χρωμοσωμάτων διπλασιάζεται και για περισσότερες επαναλήψεις οι λανθασμένες περιπτώσεις μειώνονται. Με αυτό προκύπτει το συμπέρασμα ότι για να γίνει καλύτερη εκμάθηση του μοντέλου του ΑΓΧ χρειάζονται να εκτελεστούν αρκετές επαναλήψεις.

Testing	DataSet4 with normalization 1		
Case	1	2	3
Max Generations (iterations)	100	100	100
Max Chromosomes	500	500	500
Crossover Selection Method	Roulette Wheel	Roulette Wheel	Tournament (size=100)
Crossover Method	One Point	One Point	One Point
Crossover Probability	0.9	0.1	0.1
Mutation Method	Random	Random	Random
Mutation Probability	0.01	0.01	0.01
Replacement Method	Best individuals	Best individuals	Best individuals
Best Min Fit (out of 840)	345	348	346
Wrong Normal (out of 420)	9	9	7
Wrong Pet (out of 420)	334	339	339

Πίνακας 5.6.7 Εκτελέσεις Εκμάθησης με σύνολο δεδομένων 4 με την 1^η κανονικοποίηση

Συγκρίνοντας τις περιπτώσεις 1 και 2 του πίνακα 5.6.7 παρατηρούμε ότι όταν η πιθανότητα crossover στον τερματισμό οι λάθος περιπτώσεις είναι ελάχιστα λιγότερες. Ενώ συγκρίνοντας τις περιπτώσεις 2 και 3 παρατηρούμε ότι αν εφαρμοστεί η μέθοδος tournament αντί της roulette wheel οι λάθος περιπτώσεις στο σύνολο είναι λιγότερες.

Από τις πιο πάνω εκτελέσεις φάνηκε ότι η 2^η κανονικοποίηση δίνει λιγότερες λανθασμένες περιπτώσεις. Με το σύνολο δεδομένων 5 έγινε περαιτέρω διερεύνηση στις αλλαγές των διάφορων παραμέτρων. Οι εκτελέσεις αυτές φαίνονται στους πίνακες που ακολουθούν.

Testing	DataSet 5 with normalization 2			
Case	1	2	3	4
Max Generations (iterations)	100	100	100	100
Max Chromosomes	500	500	500	500
Crossover Selection Method	Roulette Wheel	Roulette Wheel	Roulette Wheel	Roulette Wheel
Crossover Method	One Point	Two Point	One Point	Two Point
Crossover Probability	0.9	0.9	0.9	0.9
Mutation Method	Random	Random	Random	Random
Mutation Probability	0.01	0.01	0.1	0.1
Replacement Method	Best individuals	Best individuals	Best individuals	Best individuals
Best Min Fit (out of 840)	292	287	331	328
Wrong Normal (out of 420)	24	22	21	18
Wrong Pet (out of 420)	268	265	310	310

Πίνακας 5.6.8 Εκτελέσεις Εκμάθησης με σύνολο δεδομένων 5 όταν το crossover Selection Method είναι Roulette Wheel και η πιθανότητα crossover 0.9

Παρατηρείται από τον πίνακα 5.6.8 ότι όταν η μέθοδος της επιλογής για crossover είναι η Roulette Wheel η Two point crossover μέθοδος δίνει πιο λίγες περιπτώσεις λάθους. Ακόμα παρατηρούμε ότι όταν η πιθανότητα μετάλλαξης από 0.01 έγινε 0.1 οι λάθος περιπτώσεις που προέβλεψε ήταν περισσότερες.

Testing	DataSet 5 with normalization 2					
Case	1	2	3	4	5	6
Max Generations (iterations)	100	100	100	100	100	200
Max Chromosomes	500	500	1000	500	500	500
Crossover Selection Method	Roulette Wheel	Roulette Wheel	Roulette Wheel	Roulette Wheel	Roulette Wheel	Roulette Wheel
Crossover Method	One Point	Two Point	Two Point	Two Point	One Point	Two Point
Crossover Probability	0.5	0.5	0.5	0.5	0.5	0.5
Mutation Method	Random	Random	Random	Random	Random	Random
Mutation Probability	0.01	0.01	0.01	0.5	0.1	0.01
Replacement Method	Best individuals	Best individuals	Best individuals	Best individuals	Best individuals	Best individuals
Best Min Fit (out of 840)	291	277	283	387	335	285
Wrong Normal (out of 420)	26	29	24	4	16	24
Wrong Pet (out of 420)	265	248	259	383	319	261

Πίνακας 5.6.9 Εκτελέσεις Εκμάθησης με σύνολο δεδομένων 5 όταν το crossover Selection Method είναι Roulette Wheel και η πιθανότητα crossover 0.5

Όταν μειώθηκε η πιθανότητα του crossover απο 0.9 σε 0.5 σε σχέση με τις περιπτώσεις του πίνακα 5.6.8 φαίνεται ότι οι λάθος περιπτώσεις είναι πιο λίγες. Παρατηρείται ότι στις περιπτώσεις που η πιθανότητα μετάλλαξης αυξήθηκε σε 0.5 και 0.1 οι λάθος περιπτώσεις που προέβλεψε είναι αρκετά περισσότερες από άλλες περιπτώσεις. Αυτό συμβαίνει γιατί αλλάζει πολύ η λύση σε κάθε επανάληψη και να προκύπτει χειρότερη αντί καλύτερη. Από αυτές τις δοκιμές η περίπτωση 2 έδωσε τον πιο μικρό αριθμό περιπτώσεων λάθους. Όταν έγινε ξανά εκτέλεση της ίδιας περίπτωσης με 200 επαναλήψεις (περίπτωση 6) φάνηκε ότι οι περιπτώσεις λάθους αυξήθηκαν. Αυτό οφείλεται στο γεγονός ότι ο πληθυσμός των χρωμοσωμάτων αρχικοποιείται σε κάθε

διαφορετική εκτέλεση με τυχαίο τρόπο και άρα προκύπτουν διαφορετικές λύσεις κάθε φορά και άρα όταν εκτελέστηκε την πρώτη φορά να προέκυψε μια λύση καλύτερη στις 100 εκτελέσεις από ότι στις 200 που εκτελέστηκε την δεύτερη φορά.

Testing	DataSet 5 with normalization 2			
	1	2	3	4
Max Generations (iterations)	100	100	100	100
Max Chromosomes	500	500	500	500
Crossover Selection Method	Tournament (size=100)	Tournament (size=100)	Tournament (size=100)	Tournament (size=500)
Crossover Method	One Point	Two Point	Two Point	Two Point
Crossover Probability	0.9	0.9	0.9	0.9
Mutation Method	Random	Random	Random	Random
Mutation Probability	0.01	0.01	0.001	0.01
Replacement Method	Best individuals	Best individuals	Best individuals	Best individuals
Best Min Fit (out of 840)	298	309	326	312
Wrong Normal (out of 420)	15	18	16	14
Wrong Pet (out of 420)	283	291	310	298

Πίνακας 5.6.10 Εκτελέσεις Εκμάθησης με σύνολο δεδομένων 5 όταν το crossover Selection Method είναι Tournament με μέγεθος 100 και η πιθανότητα crossover 0.9

Συγκρίνοντας τις περιπτώσεις 1 και 2 του πίνακα 5.6.10 βλέπουμε ότι όταν το crossover Method είναι Tournament και η μέθοδος crossover είναι One Point οι λάθος περιπτώσεις είναι πιο λίγες. Επίσης αν συγκρίνουμε τις περιπτώσεις 2 και 3 βλέπουμε ότι η λύση που προκύπτει από τον αλγόριθμο όταν η πιθανότητα μετάλλαξης πολύ μικρή προβλέπει περισσότερες περιπτώσεις λάθος. Τέλος συγκρίνοντας τις περιπτώσεις

2 και 4 βλέπουμε ότι όταν αυξηθεί το μέγεθος του Tournament από 100 σε 500 στον ίδιο αριθμό επαναλήψεων προκύπτει λύση που προβλέπει περισσότερες περιπτώσεις λάθος.

Testing	DataSet 5 with normalization 2				
Case	1	2	3	4	5
Max Generations (iterations)	100	100	100	100	100
Max Chromosomes	500	500	500	500	500
Crossover Selection Method	Tournament (size=100)	Tournament (size=500)	Tournament (size=100)	Tournament (size=100)	Tournament (size=100)
Crossover Method	One Point	One point	Two Point	Two Point	Two Point
Crossover Probability	0.7	0.7	0.7	0.1	0.1
Mutation Method	Random	Random	Random	Random	Random
Mutation Probability	0.01	0.01	0.01	0.01	0.001
Replacement Method	Best individuals	Best individuals	Best individuals	Best individuals	Best individuals
Best Min Fit (out of 840)	303	293	280	309	356
Wrong Normal (out of 420)	18	21	22	18	11
Wrong Pet (out of 420)	285	272	258	291	345

Πίνακας 5.6.11 Εκτελέσεις Εκμάθησης με σύνολο δεδομένων 5 όταν το crossover Selection Method είναι Tournament

Συγκρίνοντας τις περιπτώσεις 1 και 2 του πίνακα 5.6.11 παρατηρούμε ότι όταν η μέθοδος crossover είναι One Point και η πιθανότητα crossover είναι 0.7 όταν αυξηθεί το μέγεθος του Tournament από 100 σε 500 προκύπτει λύση που δίνει λιγότερες περιπτώσεις λάθος. Επίσης συγκρίνοντας τις περιπτώσεις 3 και 4 προκύπτει το συμπέρασμα ότι αν μειωθεί αρκετά η πιθανότητα crossover μπορεί να προκύψει

χειρότερη λύση η οποία δίνει περισσότερες λάθος περιπτώσεις. Ενώ συγκρίνοντας τις περιπτώσεις 4 και 5 βλέπουμε ότι όταν η πιθανότητα μετάλλαξης η λύση που προκύπτει δίνει πολύ περισσότερες περιπτώσεις λάθους. Επίσης αν συγκρίνουμε τις περιπτώσεις 1 και 3 στον πίνακα 5.6.11 σε σχέση με τις περιπτώσεις 1 και 2 του πίνακα 2.6.10 παρατηρείται ότι όταν η πιθανότητα crossover ήταν μεγαλύτερη με τη μέθοδο one point crossover προέκυψε λύση που έδινε λιγότερες περιπτώσεις λάθους, κάτι αντίθετο στις περιπτώσεις 1 και 3 του 5.6.11 αφού με μικρότερη πιθανότητα crossover και με μέθοδο two point crossover προκύπτει λύση με λιγότερες περιπτώσεις λάθους.

Κεφάλαιο 6

Συμπεράσματα

Σε αυτή τη διπλωματική εργασία έγινε μια προσπάθεια εκμάθησης του μοντέλου των Ασαφών Γνωστικών Χαρτών με δύο προσεγγίσεις. Η μια προσέγγιση έκανε χρήση του γενικού γενετικού αλγορίθμου ενώ η δεύτερη ήταν μη επιβλεπόμενη μάθηση με τη χρήση του αλγορίθμου NHL.

Η εκμάθηση του μοντέλου εφαρμόστηκε για το πρόβλημα της έγκαιρης διάγνωσης του ρίσκου εμφάνισης προ-εκλαμψίας σε μια γυναίκα κατά τη διάρκεια της εγκυμοσύνης της. Με τα δεδομένα ιατρικών εξετάσεων αλλά και άλλων προσωπικών στοιχείων από περιπτώσεις γυναικών που εμφάνισαν ή όχι προ-εκλαμψία κατά τη διάρκεια της εγκυμοσύνης τους, έγινε προσπάθεια για καθορισμό του πίνακα βαρών με τη χρήση Γενετικού αλγορίθμου. Στόχος της εκμάθησης και των δύο αλγορίθμων ήταν να μάθει το μοντέλο από τις περιπτώσεις των εγκύων γυναικών και να δώσει σαν αποτέλεσμα τον βέλτιστο πίνακα των βαρών, τις σχέσεις ανάμεσα στους κρίσιμους παράγοντες που επηρεάζουν στην εμφάνιση προ-εκλαμψίας σε μια έγκυο γυναίκα. Το μοντέλο του ΑΓΧ, μετά από επιτυχημένη εκμάθηση θα ήταν ικανό να προβλέψει σωστά οποιαδήποτε νέα περίπτωση εγκύου γυναίκας και να μας πει εάν θα εμφανίσει ή όχι προ-εκλαμψία.

Δυστυχώς τα αποτελέσματα από την εκμάθηση του μοντέλου των ΑΓΧ, δεν ήταν θετικά αφού οι περιπτώσεις λάθους των εγκύων γυναικών που εμφάνισαν προ-εκλαμψία και προέβλεπε η καλύτερη λύση ήταν αρκετές σε σύγκριση με τις περιπτώσεις που δεν εμφάνισαν προ-εκλαμψία.

Παρόλα αυτά έγινε μια διερεύνηση των διαφόρων παραμέτρων και τελεστών που χρησιμοποιούνται στην εκμάθηση του μοντέλου με τη χρήση Γενετικών Αλγορίθμων. Η διαδικασία αυτή είναι αρκετά χρονοβόρα λόγω του ότι οι συνδυασμοί για τις τιμές που μπορούν να πάρουν οι παράμετροι του αλγορίθμου καθώς και οι γενετικοί τελεστές που μπορούν να χρησιμοποιηθούν είναι πάρα πολλοί. Γι αυτό το λόγο

χρειάζεται να εφαρμόσουμε τη μέθοδο δοκιμής λάθους για να προκύψει ένας συνδυασμός που να δίνει την καλύτερη λύση σε σχέση με άλλους. Αυτό που προκύπτει σαν συμπέρασμα είναι ότι για να μπορέσουμε να βρούμε την καλύτερη λύση θα πρέπει οι δοκιμές μας να εκτελούνται για ένα αρκετά μεγάλο αριθμό επαναλήψεων. Ο αριθμός των 100 επαναλήψεων για αρκετές από τις περιπτώσεις που εκτελέστηκαν δεν αποτελεί ικανοποιητικό αριθμό επαναλήψεων ώστε να εκπαιδευτεί σωστά το μοντέλο. Αν για παράδειγμα είχαμε το χρονικό περιθώριο να εκτελέσουμε τον αλγόριθμο για ένα πάρα πολύ μεγάλο αριθμό επαναλήψεων (πχ. 300 000) τότε θα μπορούσε να βελτιωθεί αρκετά η καλύτερη λύση που προκύπτει από τον αλγόριθμο. Επίσης ακόμα ένα συμπέρασμα που προκύπτει είναι ότι αν εκτελεστεί πολλές φορές ένας συγκεκριμένος συνδυασμός που θεωρείται ότι δίνει τη βέλτιστη λύση χωρίς καμία αλλαγή μπορεί να δώσει χειρότερη ή ακόμα και πολύ καλύτερη λύση. Αυτό συμβαίνει λόγω του ότι ο πίνακας όλων των λύσεων (χρωμοσώματα) σε κάθε αρχή της εκτέλεσης του αλγορίθμου δημιουργείται εντελώς τυχαία με αποτέλεσμα να δημιουργούνται χειρότερες λύσεις και να χρειάζεται περισσότερες επαναλήψεις ο αλγόριθμος για να καταλήξει στην καλύτερη λύση. Όσο αφορά τις πιθανότητες ανασυνδυασμού γονιδίων και μετάλλαξης προκύπτει το συμπέρασμα ότι καλό είναι η πιθανότητα crossover να έχει τιμή μεγαλύτερη από 0.5 και η πιθανότητα μετάλλαξης να μην είναι μεγάλη αλλά ούτε και υπερβολικά μικρή.

Θα μπορούσε σε μελλοντική εργασία ο αρχικός πίνακας με τις τιμές των βαρών στην προσέγγιση που έγινε με την χρήση γενετικών αλγορίθμων να μην αρχικοποιείται εντελώς τυχαία αφού τα βάρη (σχέσεις μεταξύ δύο παραγόντων) που μπορεί να προκύψουν να μην έχουν λογική για κάποιο γιατρό, γυναικολόγο. Θα μπορούσε να γίνει εκτέλεση με μια καθορισμένη αρχικοποίηση, ή ακόμη θα μπορούσαν να καθοριστούν όρια για συγκεκριμένα βάρη. Με αυτό τον τρόπο να καθοδηγηθεί καλύτερα η εύρεση της καλύτερης λύσης.

Η ανάπτυξη ενός μοντέλου ΑΓΧ είναι αρκετά δύσκολη και χρονοβόρα, και ειδικά για ένα θέμα που έχει να κάνει με εξειδικευμένα ιατρικά δεδομένα. Ακόμα και ο καθορισμός των κόμβων απαιτεί γνώση γύρω από το πεδίο του οποίου γίνεται προσπάθεια να μοντελοποιηθεί. Έτσι στην περίπτωση της προ-εκλαμψίας η επιλογή των πιο κρίσιμων παραγόντων που επηρεάζουν στην εμφάνιση της και στην παρούσα

διπλωματική έγινε βάση προηγούμενης βιβλιογραφίας θα μπορούσε να μελετηθεί περαιτέρω σε μελλοντική εργασία. Χρειάζεται δηλαδή να εξεταστούν κι άλλοι παράγοντες που μπορεί να επηρεάζουν στην εμφάνιση προ-εκλαμψίας σε μια έγκυο γυναίκα. Η συμβολή ενός ειδικού, γυναικολόγου ή μαιευτήρα ο οποίος έχει γνώσεις και εμπειρία με περιστατικά εγκύων γυναικών με προ-εκλαμψία θα ήταν πολύ βοηθητική για πιο ορθό καθορισμό των παραμέτρων, δηλαδή των κόμβων του μοντέλου μας.

Βιβλιογραφία

- [1] Bart Kosko, *Fuzzy Cognitive Maps*, Int. J. Man-Machine Studies (1986) 24, 65-75
- [2] C. Neocleous, M. Papaioannou, C.N. Schizas, *Fuzzy Cognitive Maps in estimating the repercussions of oil/gas exploration on politico-economic issues in Cyprus*
- [3] Chrysostomos D. Stylios, Voula C. Georgopoulos and Peter P. Groumpos, *The Use Of Fuzzy Cognitive Maps In Modeling Systems*
- [4] Costas K. Neocleous, Panagiotis Anastasopoulos, Kypros H. Nikolaides, Christos N. Schizas, Kleanthis C. Neokleous, *Neural networks to estimate the risk for preeclampsia occurrence*
- [5] Costas Neocleous, Christos Schizas, Panagiotis Anastasopoulos, Kypros Nikolaides, Kleanthis Neokleous, *Neural networks to investigate the effects of smoking and alcohol abuse on the risk for preeclampsia*
- [6] D.E. Koulouriotis, I.E. Diakoulakis, D.M. Emiris, E.N. Antonidakis and I.A. Kaliakatsos, *Efficiently Modeling And Controlling Complex Dynamic Systems Using Evolutionary Fuzzy Cognitive Maps (Invited Paper)*, International Journal of Computational Cognition (<http://www.YangSky.com/yangijcc.htm>) Volume 1, Number 2, Pages 41–65, June 2003
- [7] E.I. Papageorgiou, C.D. Stylios, P.P. Groumpos, *Active Hebbian learning algorithm to train fuzzy cognitive maps*, International Journal of Approximate Reasoning 37 (2004) 219–249
- [8] Elpiniki I. Papageorgiou, Chrysostomos Stylios, Peter P. Groumpos, *Unsupervised learning techniques for fine-tuning fuzzy cognitive map causal links*, Int. J. Human-Computer Studies 64 (2006) 727–743
- [9] Elpiniki I. Papageorgiou, Konstantinos E. Parsopoulos, Chrysostomos S. Stylios, Petros P. Groumpos, Michael N. Vrahatis, *Fuzzy Cognitive Maps Learning Using Particle Swarm Optimization*, Journal of Intelligent Information Systems, 25:1, 95–121, 2005
- [10] Elpiniki I. Papageorgiou, Νέες μέθοδοι εκμάθησης για Ασαφή Γνωστικά Δίκτυα Και Εφαρμογές στην Ιατρική και Βιομηχανία, Διδακτορική Διατριβή, 2004

- [11] Elpiniki Papageorgiou and Peter Groumpos, *A Weight Adaptation Method for Fuzzy Cognitive Maps to a Process Control Problem*, M. Bubak et al. (Eds.): ICCS 2004, LNCS 3037, pp. 515–522, 2004.
- [12] Elpiniki Papageorgiou and Peter Groumpos, Member, IEEE, *Two – stage Learning Algorithm for Fuzzy Cognitive Maps*, Second IEEE International Conference on Intelligent Systems, June 2004
- [13] Elpiniki Papageorgiou, Chrysostomos Stylios, and Peter Groumpos, *Fuzzy Cognitive Map Learning Based on Nonlinear Hebbian Rule*, T.D. Gedeon and L.C.C. Fung (Eds.): AI 2003, LNAI 2903, pp. 256-268, 2003
- [14] F. Herrera, M. Lozano & J.L. Verdegay, *Tackling RealCoded Genetic Algorithms: Operators and Tools for Behavioural Analysis*, Artificial Intelligence Review 12: 265–319, 1998
- [15] G. Büyüközkan , O. Feyzioglu, and Z. Vardaloglu, *Analyzing CPFR Supporting Factors with Fuzzy Cognitive Map Approach*, World Academy of Science, Engineering and Technology 55 2009
- [16] Guillermo A. Mendoza, Ravi Prabhu, *Participatory modeling and analysis for sustainable forest management: Overview of soft system dynamics models and applications*, Forest Policy and Economics 9 (2006) 179– 196
- [17] Heinz Muhlenbein, Dirk Schlierkamp-Voose, *Predictive Models for the Breeder Genetic Algorithm, I. Continuous Parameter Optimization*, Evolutionary Computation, Volume 1, Number 1
- [18] Karen J. Carlson, Stephanie A. Eisenstat, Terra Diane Ziporyn, *The new Harvard guide to women's health*, 2004 pages 482,483
- [19] M. Ghazanfari, S. Alizadeh, M. Fathian, D.E. Koulouriotis, *Comparing simulated annealing and genetic algorithm in learning FCM*, Applied Mathematics and Computation 192 (2007) 56–68
- [20] M. Papaioannou, *Αναφορά Π.13 – Εσωτερική αναφορά σχετικά με τις ιδιαιτερότητες παραμέτρων και αλληλεπιδράσεων*
- [21] M. Shamim Khan and Alex Chong, *Fuzzy Cognitive Map Analysis with Genetic Algorithm*
- [22] Voula C. Georgopoulos, Georgia A. Malandrakia, Chrysostomos D. Stylios, A

fuzzy cognitive map approach to differential diagnosis of specific language impairment, Artificial Intelligence in Medicine 29 (2003) 261–278

- [23] Wojciech Stach, Lukasz Kurgan, *Modeling Software Development Projects Using Fuzzy Cognitive Maps*
- [24] Wojciech Stach, Lukasz Kurgan, Witold Pedrycz, and Marek Reformat, *Evolutionary Development of Fuzzy Cognitive Maps*, The 2005 IEEE International Conference on Fuzzy Systems
- [25] Wojciech Stach, Lukasz Kurgan, Witold Pedrycz, Marek Reformat, *Genetic learning off uzzy cognitive maps*, Fuzzy Sets and Systems 153 (2005) 371–401
- [26] Zbigniew Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs*, Third Revised and Extended Edition, 1996

Διαδικτυακές Πηγές

- [27] American Pregnancy association,
<http://www.americanpregnancy.org/duringpregnancy/hcglevels.html>
- [28] Baby Center Chorionic villus sampling (CVS), Reviewed by the BabyCenter Medical Advisory Board,
<http://www.babycenter.com>
- [29] Healthier World, Τι είναι η συστολική και τι η διαστολική πίεση;, www.healthierworld.gr,
<http://www.healthierworld.gr/portal/cc7146d26842552e/31ab328e47c4ea3f/d87c68a56bc8eb80/ec69afac0d09f291/1487bea7e641dfed.html>
- [30] Introduction to Genetic Algorithms – Tutorial with interactive Java Applets,
<http://www.obitko.com/tutorials/genetic-algorithms/>
- [31] Lynne McLeod, MD, How useful is uterine artery Doppler ultrasonography in predicting pre-eclampsia and intrauterine growth restriction?, CMAJ, March 11, 2008; 178 (6). doi:10.1503/cmaj.080242,
<http://www.cmaj.ca/cgi/content/full/178/6/727>

- [32] National Center for Biotechnology Information, NCBI,
<http://www.ncbi.nlm.nih.gov/>
- [33] Peltarion Synapse
<http://www.peltarion.com/>
- [34] Preeclampsia Foundation,
<http://www.preeclampsia.org/>
- [35] U.S. National Library of Medicine, National Institutes of Health,
<http://www.nlm.nih.gov/>
- [36] Wikipedia, Fitness proportionate selection,
http://en.wikipedia.org/wiki/Fitness_proportionate_selection
- [37] WikiPedia, Tournament selection,
http://en.wikipedia.org/wiki/Tournament_selection
- [38] Άρης Γιάβρης, *Ασαφής Λογική: Ανάμεσα στο ναι και το όχι*,
<http://www.arisgiavris.gr/giavris-grammar-9.7-epikoinonia-glossa-protaseis.htm>
- [39] Γιάννης Δεβετζόγλου, *Εκλαμψία και προεκλαμψία στην εγκυμοσύνη*, Τα Νέα Online, Δημοσίευση: 9-2-2009,
<http://ygeia.tanea.gr/default.asp?pid=8&ct=13&articleID=5630&la=1>
- [40] Κέντρο Εξωσωματικής Γονιμοποίησης EmBIO, emBIO medical center,
<http://www.ivf-embryo.gr/gynaikologia/eklampsia.htm>

Παράρτημα Α

Παραθέτετε ο κώδικα στην γλώσσα προγραμματισμού MATLAB για την εκτέλεση του αλγορίθμου NHL

```
function out= NHL(Ainit,Winit,DOC)

conditionTrue=0;
size=length(Ainit);
Wnew=zeros(size,size);

cycle=0;
MAXCYCLES=1000;
g=0.98;
n=0.01;

while(conditionTrue==0)

%increasing the cycle counter
cycle=cycle+1;

%Update the weights
for j=1:size
    for i=1:size

        if((j~=i)&& Winit(j,i)~=0)
            Wnew(j,i)=g*Winit(j,i)+n*Ainit(j)*(Ainit(i)-
Ainit(j)*Winit(j,i));
        else
            Wnew(j,i)=Winit(j,i);
        end
    end
end

%Calculate Anew
Anew=sigmoidFunction(Ainit+Ainit*Wnew);

val=termFunc2(Ainit,Anew,DOC);

    if(cycle==MAXCYCLES || val==1)
        conditionTrue=1;
    end

Ainit=Anew;
Winit=Wnew;
out=Wnew;

disp('WNEW');
disp(Wnew);
disp('ANEW');
disp(Anew);
```

```

end

end

function output = termFunc2(Ak,Ak1,DOC)

count=0;
numDOC=length(DOC);
e=0.002;

for i=1:numDOC
    if(abs(Ak1(DOC(i)) - Ak(DOC(i))) < e)
        count=count+1;
    end
end

if(count==numDOC)
    output=1;
else
    output=0;
end

end

function outSigm = sigmoidFunction(x)

outSigm= 1./(1+exp(-x));

end

```

Παράρτημα Β

Πίνακας διαφόρων παραγόντων που πιθανόν να επηρεάζουν στην εμφάνιση προ-εκλαμψίας σε μια έγκυο γυναίκα. Η πρώτη στήλη είναι το όνομα του παράγοντα, ενώ στη δεύτερη στήλη δίνεται μια μικρή διευκρίνιση.

Όνομα παράγοντα	Περιγραφή
GA-days	Gestation age at the time of scan in days: Η ηλικία του εμβρύου σε ημέρες τη στιγμή που έγινε η εξέταση υπέρηχου
GA-weeks	Gestation age at the time of scan in weeks: Η ηλικία του εμβρύου σε εβδομάδες τη στιγμή που έγινε η εξέταση του υπέρηχου. Ο υπολογισμός του γίνεται με μια απλή διαίρεση της ηλικίας σε ημέρες δια τις εφτά ημέρες της εβδομάδας.
Age	Η ηλικία της εγκύου μητέρας. Το πεδίο των ηλικιών κυμαίνεται από την μικρότερη ηλικία των 15, 16 μέχρι 50.
Ethnicity	Εθνικότητα της εγκύου μητέρας. Οι κατηγορίες είναι: 1: White, 2: Black, 3: Indian, 4: Chinese, 5: Mixed
Wt-now	Weight at the time of scan: Το βάρος της εγκύου μητέρας τη στιγμή που έγινε η εξέταση του υπέρηχου.
Height	Το ύψος της εγκύου μητέρας, σε εκατοστόμετρα (cm)
BMI – now	Body mass index at the time of scan: Δείκτης μάζας σώματος, ιατρική ένδειξη για τον υπολογισμό βαθμού παχυσαρκίας. Υπολογίζεται από τον τύπο: (βάρος σε κιλά)/(Ύψος σε μέτρα) ²
Smoking	Κάπνισμα κατά την εγκυμοσύνη. 1: καπνιστής 0: μη καπνιστής

Alcohol	Κατανάλωση αλκοόλ από την έγκυο μητέρα κατά τη διάρκεια της εγκυμοσύνης. 1: κατανάλωνε αλκοόλ 0: δεν κατανάλωνε αλκοόλ
DrugAbuse	Χρήση ναρκωτικών ουσιών από την έγκυο μητέρα κατά τη διάρκεια της εγκυμοσύνης. 1: χρήστης ναρκωτικών ουσιών 0: δεν ήταν χρήστης ναρκωτικών ουσιών
FH-PET-mother	Ο παράγοντας ορίζει κατά πόσο η μητέρα της εγκύου μητέρας είχε εμφανίσει προ-εκλαμψία. 1: είχε εμφανίσει προ-εκλαμψία 0: δεν είχε εμφανίσει προ-εκλαμψία
FH-PET-sister	Ο παράγοντας ορίζει κατά πόσο η αδερφή της εγκύου μητέρας είχε εμφανίσει προ-εκλαμψία. 1: είχε εμφανίσει προ-εκλαμψία 0: δεν είχε εμφανίσει προ-εκλαμψία
Para	Parity Πόσες εγκυμοσύνες είχε προηγουμένως η έγκυος μητέρα. 1: Multip: Δεν είναι η πρώτη της εγκυμοσύνη. 0: Nullip: Πρώτη της εγκυμοσύνη.
Prev. PET	Ο παράγοντας αυτός ορίζει αν η έγκυος μητέρα είχε εμφανίσει στο ιστορικό της προ-εκλαμψία. 1: Έχει εμφανίσει στο ιστορικό της προ-εκλαμψία 0: Δεν έχει εμφανίσει προ-εκλαμψία.
Conception	Σύλληψη, ο τρόπος με τον οποίο η έγκυος γυναίκα έμεινε έγκυος. 1: Spontaneous: δηλαδή με φυσιολογική σεξουαλική επαφή 2: IVF: intro vitro fertilization, η γονιμοποίηση γίνεται έξω από το ανθρώπινο σώμα, παιδί του σωλήνα. 3: Ovulation: Όταν η έγκυος γυναίκα παίρνει φάρμακα για να προκληθεί ωορρηξία.
Medical pre-pregnancy	Ο παράγοντας αυτός περιγράφει την ιατρική κατάσταση της εγκύου γυναίκας πριν την εγκυμοσύνη. Δηλαδή αν η έγκυος γυναίκα υπέφερε από

	κάποια ασθένεια όπως: 1: Chronic hypertension, 2: Antiphospholid syndrome, 3: Asthma 4: Crohns disease, 5: Depression, 6: Sickle cell disease 7: Diabetes (IDDM), Epilepsy, Hypothyroidism, Hyperthyroidism, Infection: HIV ή 0: No medical cond group (η έγκυος γυναίκα δεν ανήκει σε κάποια από τις πιο πάνω ιατρικές ομάδες, και άρα δεν υποφέρει από κάποια ασθένεια)
Medical Pregnancy	Ο παράγοντας αυτός αναφέρει αν η έγκυος γυναίκα είχε κάποιο ιατρικό πρόβλημα, εκτός προ-εκλαμψίας, κατά τη διάρκεια της εγκυμοσύνης της, δηλαδή: 1: Cholestasis 2: Gestational Diabetes
Drugs	Ο παράγοντας αυτός αναφέρει αν η έγκυος γυναίκα έπαιρνε κάποιου είδους φάρμακα κατά τη διάρκεια της εγκυμοσύνης της, όπως: 1: Antihypertensives, 2: Insulin, 3: Steroids, 4: Betamimetic 5: Combined asthma, 6: Thyroxine, 7: Aspirin, heparin, 8: Antithyroid, 9: Antiepileptics, 10: Lithium, 11: Antidepressant 12: Antiretroviral , 15: None
CRL	Crown rump length: Το μήκος του μωρού από την κορυφή μέχρι τους γλουτούς. Το μήκος του μωρού κυμαίνεται από 45 μέχρι 84mm
NT	Nuchal translucency: Αυχενική Διαφάνεια: Μέτρηση του πάχους του υγρού στον αυχένα του βρέφους σε mm
PAPP	Πρωτεΐνη PAPP: Pregnancy Associated plasma protein – A Πρωτεΐνη που παράγεται από τον πλακούντα και το τοίχωμα της μήτρας κατά τη διάρκεια της εγκυμοσύνης. Μετρείται σε μονάδες ανά λίτρο (Units per Litre).
bhCG	Beta human chorionic gonadotropin: είναι μια ορμόνη που παράγεται κατά την ανάπτυξη της εγκυμοσύνης και στη συνέχεια παράγεται σε μεγάλες ποσότητες από τον πλακούντα.
PAPP MoM	Corrected PAPP-A expressed as a multiple of median MoM

bhCG-MoM	Corrected bhCG expressed as a multiple of median MoM
CVS	Chorionic Villus Sampling Είναι μια εξέταση, όπως και η αμνιοκέντηση η οποία όμως μπορεί να γίνει στα αρχικά στάδια της εγκυμοσύνης (στις 10 με 12 εβδομάδες) αντί στις 15 με 20 που γίνεται η εξέταση της αμνιοκέντησης. Κατά την εξέταση αυτή γίνεται εξέτασης του πλακούντα για καθορισμό γενετικής ανωμαλίας ή ανωμαλίας σε χρωμόσωμα του εμβρύου [28] 1: Amnio: Η έγκυος γυναίκα έχει κάνει την εξέταση της αμνιοκέντησης 2: CVS: Η έγκυος γυναίκα έχει κάνει την εξέταση Chorionic Villus Sampling 0: Η έγκυος γυναίκα δεν έχει κάνει καμία από τις δύο εξετάσεις
Karyo	Ο παράγοντας αυτός δείχνει το αποτέλεσμα του Καρυότυπου
Other defects	Ο παράγοντας αυτός αναφέρει αν η έγκυος γυναίκα είχε οποιαδήποτε άλλη ανωμαλία
Ut mPI	Uterine artery pulsatility index Υπολογίζεται από τις κυματομορφές του υπερηχογραφήματος Doppler. Είναι ένας δείκτης της ταχύτητας της ροής του αίματος σε αρτηρίες της μήτρας. [31] Mean between left (Left Ut PI) and right (Right Ut PI)
Left Ut PI	Uterine pulsatility index αριστερής αρτηρίας στη μήτρα
Right Ut PI	Uterine pulsatility index δεξιάς αρτηρίας στη μήτρα
Systolic BP	Συστολική πίεση καλείται η πίεση που ασκεί το αίμα στις αρτηρίες όταν η καρδιά συσπάται για να προωθήσει το αίμα μέσω των αρτηριών προς τα όργανα του σώματος [29]
Diastolic BP	Diastolic blood pressure – Διαστολική πίεση καλείται η πίεση που ασκεί το αίμα στις αρτηρίες όταν η καρδιά χαλαρώνει, ανάμεσα στους κτύπους της, για να δεχθεί νέο αίμα [29]

MAP	MAP- Μέση Αρτηριακή Πίεση Υπολογίζεται ως εξής: $MAP = DBP + ((SBP - DBP) / 3) = (2 \times DBP + SBP) / 3$
BP. no	Πόσες μετρήσεις αρτηριακής πίεσης έχουν ληφθεί
Card - Pulse	Καρδιακός παλμός της εγκύου μητέρας
Card - SV	Card Stroke volume - είναι ο όγκος του αίματος που αντλείται από την μια κοιλία της καρδιάς στην άλλη σε κάθε κτύπημα.
Card-Index SV	Η τιμή του Card – SV διαιρεμένη με την επιφάνεια του σώματος
Card-CO	Cardiac Output είναι ο όγκος του αίματος που αντλείται από την καρδιά, συγκεκριμένα από την αριστερή και δεξιά κοιλία της καρδιάς στο χρονικό διάστημα του ενός λεπτού. Μετρίεται σε λίτρα ανά λεπτό.
Card-Index CO	Η τιμή του Cardiac Output διαιρεμένη με την επιφάνεια του σώματος
Ur-Alb mg/L	Urine – Albumin (mg/L) Πόσα μίλιγραμμάρια(mg) πρωτεΐνης albumin περιέχονται σε 1litre ούρων Ο όρος albumin αναφέρεται σε οποιαδήποτε πρωτεΐνη η οποία έχει την ικανότητα να διαλύεται στο νερό. [32]
Ur-Creat mmol/L	Urine – Creatinine (mmol/L) Πόσα μίλιγραμμομόρια κρεατινίνης περιέχονται σε 1 λίτρο ούρων. Η κρεατινίνη σχηματίζεται από την κρεατίνη, που είναι ένα σημαντικό στοιχείο στους μυς. Αφαιρείται ολοκληρωτικά από το σώμα μέσω των νεφρών [32]
Ur-ACR	Είναι η αναλογία albumin/creatinine
Outcome	Το αποτέλεσμα της κυοφορίας της εγκύου γυναίκας. -Θάνατος του εμβρύου μέσα στην μήτρα

	-LB-NND – Το έμβρυο γεννήθηκε ζωντανό αλλά πέθανε εντός μιας εβδομάδας -Αποβολή εμβρύου -Το έμβρυο γεννήθηκε ζωντανό -TOP – Τερματισμός εγκυμοσύνης
Indication TOP	Ο παράγοντας αυτός εξηγεί στην περίπτωση που το αποτέλεσμα της κυοφορίας ήταν ο τερματισμός, την ένδειξη για την οποία τερματίστηκε η εγκυμοσύνη
Out-GA(d)	Η ηλικία του εμβρύου την στιγμή που γεννήθηκε, σε ημέρες
Sex	Το φύλο του εμβρύου 1: άρρεν 2: θήλυ
BW	Το βάρος του εμβρύου την στιγμή που γεννήθηκε, σε γραμμάρια
Labour	Ο τρόπος με τον οποίο γεννήθηκε το μωρό, αν γεννήθηκε αυθόρμητα ή προκάλεσαν την γέννηση του για παράδειγμα με καισαρική τομή
Indication labour	Ο παράγοντας αυτός αναφέρει για τις ενδείξεις τις οποίες οδήγησαν στο να προκαλέσουν την γέννηση του βρέφους
Delivery	Ο τρόπος με τον οποίο γεννήθηκε το μωρό, για παράδειγμα αν γεννήθηκε με κολπική εξαγωγή ή έγινε καισαρική τομή
Indication CS	Ο παράγοντας αυτός αναφέρει την ένδειξη για την οποία η γέννα έγινε με καισαρική τομή
PET	Ο παράγοντας αυτός αναφέρει σε ποια κατηγορία ανήκει η έγκυος γυναίκα, δηλαδή αν είχε preeclampsia, χρόνια υπέρταση κλπ. 1: PET (preeclampsia) 2: PIH (pregnancy induces hypertension) 3: Chronic hypertension 4: PET - Other 5: HELLP syndrome

	0: Normal
--	-----------

Παράρτημα Γ

Εισάγεται το θέμα που προτίθεστε να διατυπώσετε στο παράρτημα αυτό. Εισάγεται το θέμα που προτίθεστε να διατυπώσετε στο παράρτημα αυτό. Εισάγεται το θέμα που προτίθεστε να διατυπώσετε στο παράρτημα αυτό.

