

Ατομική Διπλωματική Εργασία

**ΣΥΣΤΗΜΑΤΑ ΥΠΟΛΟΓΙΣΤΙΚΗΣ ΝΟΗΜΟΣΥΝΗΣ
ΓΙΑ ΠΡΟΒΛΕΨΗ ΧΡΩΜΟΣΩΜΙΚΩΝ ΑΝΩΜΑΛΙΩΝ ΤΟΥ
ΕΜΒΡΥΟΥ ΚΑΤΑ ΤΗΝ ΕΚΓΥΜΟΣΥΝΗ**

Χαράλαμπος Χρίστου

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΥΠΡΟΥ



ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

Μάιος 2013

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΥΠΡΟΥ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

ΣΥΣΤΗΜΑΤΑ ΥΠΟΛΟΓΙΣΤΙΚΗΣ ΝΟΗΜΟΣΥΝΗΣ
ΓΙΑ ΠΡΟΒΛΕΨΗ ΧΡΩΜΟΣΩΜΙΚΩΝ ΑΝΩΜΑΛΙΩΝ ΤΟΥ ΕΜΒΡΥΟΥ ΚΑΤΑ ΤΗΝ
ΕΚΓΥΜΟΣΥΝΗ

Χαράλαμπος Χρίστου

Επιβλέπων Καθηγητής
Χρίστος Ν. Σχίζας

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

Μάιος 2013

Ευχαριστίες

Θέλω να ευχαριστήσω ιδιαίτερα τον επιβλέποντα καθηγητή μου, κ. Χρίστο Ν. Σχίζα για την εμπιστοσύνη που μου έδειξε με την ανάθεση της παρούσας διπλωματικής εργασίας, αλλά και για την στήριξη που μου παρείχε και τις οδηγίες ούτως ώστε να καταστεί δυνατή η ολοκλήρωση της.

Ευχαριστώ επίσης όλους του συμφοιτητές μου, που μου στάθηκαν και με στήριξαν σε όλες τις τυχόν δυσκολίες ή τα προβλήματα που είχα συναντήσει εντός και εκτός Πανεπιστημίου αυτά τα χρόνια.

Τέλος, ευχαριστώ την οικογένεια μου για την πίστη τους προς εμένα, την οικονομική και ψυχολογική υποστήριξη και τη συμπαράστασή τους κατά τη διάρκεια των σπουδών μου.

Περίληψη

Σκοπός της εργασίας αυτής είναι η δημιουργία υπολογιστικών μοντέλων με χρήση τεχνητών νευρωνικών δικτύων, για να εκπαιδευτούν ούτως ώστε να είναι σε θέση να προβλέπουν κάποιες συγκεκριμένες χρωμοσωμικές ανωμαλίες του εμβρύου κατά τη διάρκεια της πρώτης τριμηνίας της εγκυμοσύνης. Οι χρωμοσωμικές ανωμαλίες που μελετήθηκαν είναι το σύνδρομο Down. Η αρχική βάση δεδομένων που χρησιμοποιήσα περιείχε 29791 περιπτώσεις φυσιολογικών περιστατικών, 49 περιπτώσεις συνδρόμου T13, 154 περιπτώσεις συνδρόμου T18, 380 περιπτώσεις συνδρόμου T21, 34 περιπτώσεις συνδρόμου Triploidy και 74 περιπτώσεις συνδρόμου Turner. Στη βάση δεδομένων που χρησιμοποιήσα στα μοντέλα νευρωνικών δικτύων επιλέχθηκαν 124 περιστατικά για την τρισωμία 21 και 29791 φυσιολογικά περιστατικά, 29915 περιστατικά στο σύνολο.

Η εργασία βασίστηκε στην κατασκευή ενός αλγόριθμου στη γλώσσα MATLAB που αρχικά χωρίζει τυχαία το σύνολο των δεδομένων σε δεδομένα εκπαίδευσης και επαλήθευσης, σύμφωνα με ένα αριθμό Fold Validation που δίνει ο χρήστης. Επιπρόσθετα, δημιουργεί μοντέλα αλλάζοντας τις παραμέτρους και τις αρχιτεκτονικές των νευρωνικών δικτύων, σύμφωνα πάντα με οδηγίες που του δίνει ο χρήστης. Επίσης, ο αλγόριθμος υποστηρίζει μια τεχνική η οποία αγνοεί ένα υποσύνολο των χαρακτηριστικών εισόδου και εκπαιδεύει το δίκτυο χωρίς αυτά, ούτως ώστε να εξεταστούν πόσο καλές προβλέψεις μπορούμε να έχουμε, αγνοώντας κάποια χαρακτηριστικά ή κάποιο συνδυασμό χαρακτηριστικών. Στη συνέχεια ο αλγόριθμος υπολογίζει τα ποσοστά επιτυχίας για κάθε κατηγορία της βάσης δεδομένων και παρουσιάζει τα αποτελέσματα, τα οποία μελέτησα και σχολίασα.

Τέλος, παρουσιάζονται κάποια συμπεράσματα που εξήχθησαν μελετώντας τα αποτελέσματα των πειραμάτων, αλλά γίνεται και μια σύγκριση κάποιων συμπερασμάτων από προηγούμενες μελέτες με τα δικά μου συμπεράσματα και αποτελέσματα.

Περιεχόμενα

Κεφάλαιο 1	Εισαγωγή.....	1
	1.1 Νοημοσύνη	1
	1.2 Κλασσική Τεχνητή Νοημοσύνη	2
	1.3 Υπολογιστική Νοημοσύνη	3
Κεφάλαιο 2	Τεχνικές Υπολογιστικής Νοημοσύνης.....	5
	2.1 Τεχνητά Νευρωνικά Δίκτυα	5
	2.2 Εξελικτικά Συστήματα	10
	2.3 Συστήματα Ασαφούς Λογικής	12
Κεφάλαιο 3	Προηγούμενες Μελέτες.....	14
	3.1 Διπλωματικές εργασίες στο Πανεπιστήμιο Κύπρου	14
	3.2 Εργασίες στο Πανεπιστήμιο Κύπρου	15
	3.3 Εργασίες σε διεθνές επίπεδο	16
Κεφάλαιο 4	Ανάλυση θέματος και εξήγηση.....	19
	4.1 Ορισμός προβλήματος	19
	4.2 Ερευνητική ερώτηση	23
	4.3 Ανάλυση χαρακτηριστικών	23
	4.4 Παραδοσιακές μέθοδοι αντιμετώπισης	25
	4.5 Σύγκριση και συζήτηση μεθόδων	26
Κεφάλαιο 5	Πρόταση Εργασίας.....	28
	5.1 Συνδυασμός Τεχνικών	28
	5.2 Τεχνική 1	30
	5.3 Τεχνική 2	31
	5.4 Τεχνική 3	32
	5.5 Αποτελέσματα	33

ΚΕΦΑΛΑΙΟ 6	Συμπεράσματα	53
	6.1 Συμπεράσματα	53
ΚΕΦΑΛΑΙΟ 7	Θέματα Μελλοντικής Μελέτης.....	59
	7.1 Θέματα μελλοντικής μελέτης	59
	Αναφορές	60
Π α ρ α ρ τ η μ α Α.....		A-1

Κεφάλαιο 1

Εισαγωγή

1.1 Νοημοσύνη

1.2 Κλασσική Τεχνητή Νοημοσύνη

1.3 Υπολογιστική Νοημοσύνη

1.1 Νοημοσύνη

Η γεφύρωση του χάσματος ανάμεσα στον άνθρωπο και στις μηχανές αποτελεί το μεγαλύτερο στόχο της Επιστήμης των Υπολογιστών. Γύρω στο 1950 τέθηκε το ερώτημα κατά πόσον οι υπολογιστές είναι τόσο ευφυείς, ώστε να μπορούν να μιμούνται τις ανθρώπινες συμπεριφορές. Η ανάγκη για επεξεργασία τεράστιας ποσότητας δεδομένων για γρήγορα και αξιόπιστα συμπεράσματα μας οδήγησε στην προσπάθεια της μοντελοποίησης των λειτουργιών του ανθρώπινου τρόπου σκέψης. Ακόμη και σήμερα, αν και η επιστήμη αυτή έχει εμφανίσει αλματώδη ανάπτυξη, δεν έχει καταφέρει να παρουσιάσει ηλεκτρονικούς υπολογιστές απαλλαγμένους από τα κλασσικά μειονεκτήματα μιας μηχανής, όπως ο τρόπος σκέψης, τα περιθώρια εκπαίδευσης και εξέλιξης, τα οποία είναι πολύ περιορισμένα και εξαρτώμενα από το στάδιο που βρίσκεται η τεχνολογία. Ο τομέας που ασχολείται με τη μοντελοποίηση των λειτουργιών του ανθρώπινου τρόπου σκέψης είναι η Υπολογιστική Νοημοσύνη, η οποία αποτελεί κομμάτι της Τεχνητής Νοημοσύνης.

Νοημοσύνη είναι γενικά η ιδιότητα κάποιας οντότητας να μπορεί να αναγνωρίζει και να αντιμετωπίζει καταστάσεις, μέσω κάποιων εμπειριών της. Η ικανότητα, δηλαδή, να εκπαιδεύεται κάποιος από παρόμοιες καταστάσεις του παρελθόντος, να βγάζει συμπεράσματα και να χρησιμοποιεί τη λογική ούτως ώστε να είναι σε θέση να λάβει κάποιες

αποφάσεις όταν βρει μπροστά του καινούρια δεδομένα. Η νοημοσύνη για τον άνθρωπο είναι συνδυασμός γενετικού υλικού και πολύχρονης εγκεφαλικής δραστηριότητας, δηλαδή κάποια είδη νοημοσύνης ο κάθε άνθρωπος τα έχει από τη φύση του υψηλότερα, ενώ βελτιώνονται όταν ο άνθρωπος αποκτά εμπειρίες και γνώσεις. Η νοημοσύνη δεν καθορίζεται από ένα μόνο χαρακτηριστικό, αλλά από πολλά, τα οποία όλα μαζί την προσδιορίζουν. Ένας άλλος ορισμός της, είναι η γενική ικανότητα που εκφράζεται μέσα από τις διαδικασίες υπολογισμών, συλλογισμών, λογικής, διακρίβωσης, μάθησης, χρήσης γλώσσας, αντίληψης του περιβάλλοντος σε διαφορετικούς βαθμούς λεπτομέρειας, εξοικείωσης σε νέο περιβάλλον, αυτοδιόρθωσης, και επινόησης.

Παράδειγμα νοημοσύνης είναι η ικανότητα του ανθρώπου να μαθαίνει να κάνει δύσκολες μαθηματικές πράξεις με το μυαλό του. Μερικά είδη της νοημοσύνης είναι η λεκτική, η σωματική, η χωρική, η μουσική και η αριθμητική.

1.2 Κλασσική Τεχνητή Νοημοσύνη

Πιο κάτω βλέπουμε σε συντομία ποιοι κλάδοι και πως συνεισέφεραν ιδέες και τεχνικές της Τεχνητής Νοημοσύνης. Από το 400 π.Χ. διάφοροι φιλόσοφοι αντιλήφθηκαν την Τεχνητή Νοημοσύνη εξετάζοντας την ιδέα ότι η νόηση μοιάζει με μηχανή, ότι λειτουργεί βασιζόμενη σε γνώσεις κωδικοποιημένες και ότι η σκέψη χρησιμοποιείται για να επιλέγεται η ενέργεια που θα χρησιμοποιηθεί. Πιο πρόσφατα ο κλάδος των μαθηματικών παρείχε τα εργαλεία για το χειρισμό προτάσεων λογικής βεβαιότητας καθώς και αβέβαιων πιθανοτικών. Ο κλάδος της ψυχολογίας ισχυρίστηκε ότι οι άνθρωποι και τα ζώα μπορούν να θεωρηθούν μηχανές επεξεργασίας πληροφοριών. Τέλος καταλήγουμε στους επιστήμονες των υπολογιστών οι οποίοι έκαναν δυνατή την εφαρμογή της Τεχνητής Νοημοσύνης η οποία έκανε γρηγορότερες προόδους την τελευταία δεκαετία λόγω και της μεγαλύτερης χρήσης της επιστημονικής μεθόδου στον πειραματισμό, αλλά και την παραγωγή υπολογιστών με μεγάλες δυνατότητες υπολογισμού και μνήμης.

Η αρχή της θεωρίας που υποστηρίζει ότι η νόηση είναι προϊόν ενός λειτουργικού υπολογιστικού μηχανισμού μπορεί να αποδοθεί στον πατέρα της Τεχνητής Νοημοσύνης, τον άγγλο μαθηματικό και θεωρητικό της λογικής, Alan Turing. Το 1950, ο Turing δημοσίευσε ένα νοητικό πείραμα που το ονόμασε παιχνίδι της μίμησης με το οποίο ήθελε να δείξει, ότι

όπως σε ένα παιχνίδι μεταξύ ανθρώπων οι παίκτες σκέφτονται και πράττουν, έτσι και ένας υπολογιστής που μπορεί να παίζει αποτελεσματικά σε αυτό το παιχνίδι, σκέφτεται και πράττει σαν ένα ον με ανθρώπινη νοημοσύνη.

Η κλασσική Τεχνητή Νοημοσύνη χρησιμοποιούσε συμβολική αναπαράσταση της γνώσης και τεχνικές ανάλυσης υψηλού επιπέδου. Η προσέγγιση αυτή είχε πολλούς περιορισμούς και έφτασε σε ένα σημείο που δεν πρόσφερε τίποτε, έτσι η έρευνα στράφηκε σε τεχνικές ανάλυσης χαμηλού επιπέδου. Η βασική ιδέα ήταν αντί να αναπαριστούμε τη γνώση σε ένα υψηλό επίπεδο, αντικαθιστούμε τα μοντέλα αυτά με αλγορίθμους που στηρίζονται στην αριθμητική επεξεργασία της πληροφορίας, δηλαδή υπολογιστικά. Έτσι καταλήξαμε στην Υπολογιστική Νοημοσύνη η οποία υποδηλώνει κάθε είδος νοημοσύνης που μπορεί να εκδηλωθεί με υπολογιστικές διαδικασίες.

1.3 Υπολογιστική Νοημοσύνη

Η Υπολογιστική Νοημοσύνη είναι ο τομέας της επιστήμης των υπολογιστών που ασχολείται με τη σχεδίαση και την υλοποίηση προγραμμάτων τα οποία ακολουθούν τις ανθρώπινες γνωστικές ικανότητες, έχοντας έτσι χαρακτηριστικά που μοιάζουν με ανθρώπινη συμπεριφορά, όπως η επίλυση προβλημάτων, η αντίληψη μέσω της όρασης, η μάθηση και η εξαγωγή συμπερασμάτων, δίνοντας την εντύπωση ότι η μηχανή είναι άνθρωπος και μπορούμε να του συμπεριφερθούμε ανάλογα, να μιλήσουμε μαζί του, να μάθουμε πράγματα ή ακόμα και να παίζουμε. Η υπολογιστική τεχνητή νοημοσύνη βασίζεται στη μάθηση μέσω επαναληπτικών διαδικασιών. Οι τρεις βασικοί μέθοδοι της υπολογιστικής νοημοσύνης είναι οι εξής:

- Τα Τεχνητά Νευρωνικά Δίκτυα
- Η Ασαφής Λογική
- Η Εξελικτική Υπολογιστική

Εφαρμογές της Υπολογιστικής Νοημοσύνης συναντούμε στην Ιατρική, στη Βιολογία, στο Στρατό, στη Βιομηχανία και σε προβλέψεις για το Περιβάλλον. Μερικά παραδείγματα από

τις ποιο πάνω περιοχές είναι η διάγνωση ιατρικών παθήσεων, η ανάγνωση και ανάλυση των ακτινών X, η πρόβλεψη καιρού, η ανάλυση επικινδυνότητας δανείων, η αναγνώριση πλαστογραφιών και η πρόβλεψη ζήτησης αγαθών.

Σήμερα, γίνεται μεγάλη προσπάθεια για την ανάπτυξη τεχνικών που προβλέπουν τις φυσικές καταστροφές, προκειμένου να ληφθούν προληπτικά μέτρα. Σε αυτή την εργασία [1], η εξαγωγή των ποσοτικών κανόνων συσχέτισης (QAR) χρησιμοποιείται για την ανακάλυψη του μοντέλου συμπεριφοράς των σεισμών, ούτως ώστε να προβλέψουμε τους επόμενους. Το σύνολο των δεδομένων που χρησιμοποιείται σε αυτή την εργασία έχει ανακτηθεί από τον κατάλογο του Ισπανικού Γεωγραφικού Ινστιτούτου, (SGI), το οποίο περιέχει τη θέση και το μέγεθος του σεισμού. Επιπλέον, η παράμετρος b-value [9] του νόμου του Gutenberg-Richter έχει υπολογιστεί, καθώς αντανακλά την τεκτονική και τις γεωφυσικές ιδιότητες των πετρωμάτων, καθώς και τις μεταβολές της πίεσης υγρού στην επιφάνεια [10]. Έτσι, κάθε δείγμα αποτελείται από τέσσερα χαρακτηριστικά: Μέγεθος τρέχων σεισμού, ώρα κατά την οποία σημειώθηκε ο σεισμός, την τιμή b-value, και το μέγεθος του προηγούμενου σεισμού. Σεισμοί με μέγεθος μικρότερο από 3.0 έχουν αφαιρεθεί από το σύνολο δεδομένων. Έχουν αφαιρεθεί επίσης οι μετασεισμικές δονήσεις για να αποφευχθούν εξαρτήσεις μεταξύ των δεδομένων. Τα δεδομένα περιέχουν 873 σεισμούς από το 1981 μέχρι το 2008. Η μελέτη αυτή είχε επιτυχία της τάξεως του 83,0%.

Σε αυτή τη μελέτη [3] γίνεται προσπάθεια για πρόβλεψη της επιχειρηματικής αποτυχίας (πτώχευσης) μέσω τεχνικών Υπολογιστικής Νοημοσύνης. Πρόκειται για μια μελέτη που παρουσιάζει την τεχνική του γραμμικού γενετικού προγραμματισμού (LGP) ως την πιο αποδοτική τεχνική για την κατασκευή μοντέλων πρόβλεψης για αυτό το πρόβλημα. Η εκπαίδευση γίνεται με 140 δείγματα πτωχευμένων και μη πτωχευμένων επιχειρήσεων του Ιράν. Συγκεκριμένα τα δεδομένα που χρησιμοποιήθηκαν για την ανάπτυξη του μοντέλου αποτελούνται από οικονομικά δεδομένα 68 πτωχευμένων και 72 μη πτωχευμένων επιχειρήσεων από το Ιράν, από το 1999 μέχρι το 2006. Τα δεδομένα ελήφθησαν από το Χρηματιστήριο της Τεχεράνης. Τα χαρακτηριστικά των δεδομένων που λήφθηκαν υπόψη είναι: Πωλήσεις σε περιουσιακά στοιχεία, Λειτουργικά έσοδα από τις πωλήσεις, Ταχεία περιουσιακά στοιχεία προς συνολικά στοιχεία ενεργητικού και Συνολική ευθύνη στο σύνολο του ενεργητικού. Τα αποτελέσματα δείχνουν ότι τα ποσοστά επιτυχίας της μελέτης αυτής φτάνουν το 99% στα δεδομένα εκπαίδευσης και 95% στα δεδομένα επαλήθευσης.

Κεφάλαιο 2

Τεχνικές Υπολογιστικής Νοημοσύνης

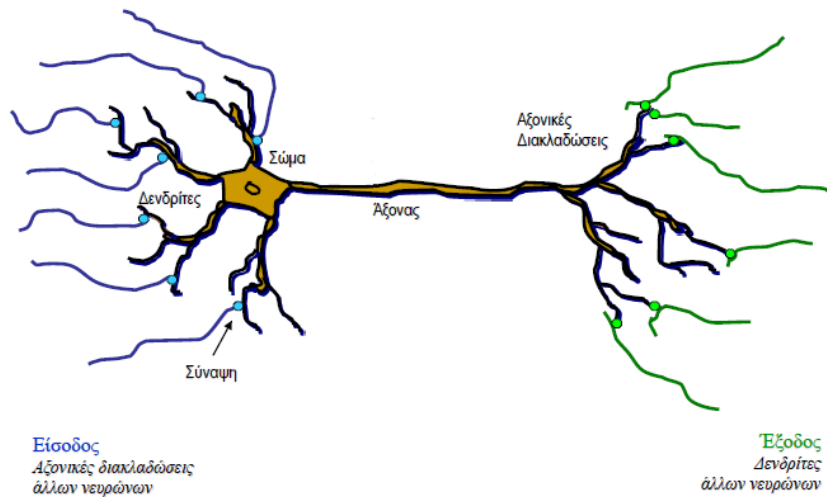
2.1 Τεχνητά Νευρωνικά Δίκτυα

2.2 Εξελικτικά Συστήματα

2.3 Συστήματα Ασαφούς Λογικής

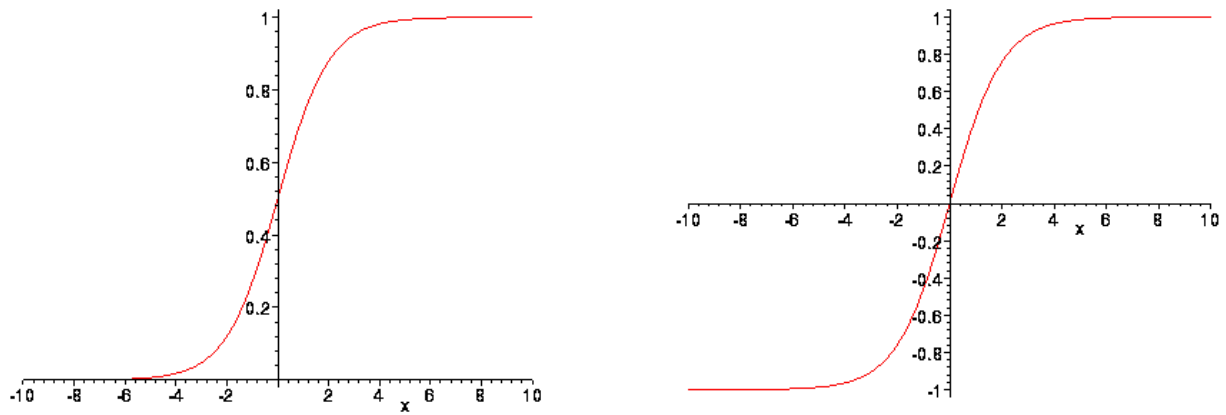
2.1 Τεχνητά Νευρωνικά Δίκτυα

Τα νευρωνικά δίκτυα αναπτύχθηκαν μέσα από έρευνες της Τεχνητής Νοημοσύνης τις τελευταίες δεκαετίες. Οι έρευνες αυτές βασίζονται στη δυνατότητα εξόρυξης γνώσης μέσα από βιολογικά νευρωνικά συστήματα, προσπαθώντας να μοντελοποιήσουν τη δομή των χαμηλών επιπέδων του εγκεφάλου. Ο εγκέφαλος αποτελείται από περίπου 10 δισεκατομμύρια βιολογικούς νευρώνες, οι οποίοι είναι μαζικά διασυνδεδεμένοι μεταξύ τους με πολλές χιλιάδες διασυνδέσεις ανά νευρώνα και επικοινωνούν μεταξύ τους μέσω συνάψεων. Ο νευρώνας έχει μια διακλαδωτική διάρθρωση εισόδων, τους δενδρίτες, και μια διακλαδωτική δομή εξόδων, τον άξονα. Από τους δενδρίτες μεταφέρεται ένα ηλεκτροχημικό σήμα προς τον νευρώνα που καταλήγουν. Όταν το συνολικό σήμα το οποίο λαμβάνεται από τους δενδρίτες, υπερβεί ένα συγκεκριμένο επίπεδο, δηλαδή, το κατώτατο όριο βολής (firing threshold), τότε ο άξονας ενεργοποιείται και πυροδοτεί ένα ηλεκτροχημικό σήμα κατά μήκος του (σχήμα 2.1.1).

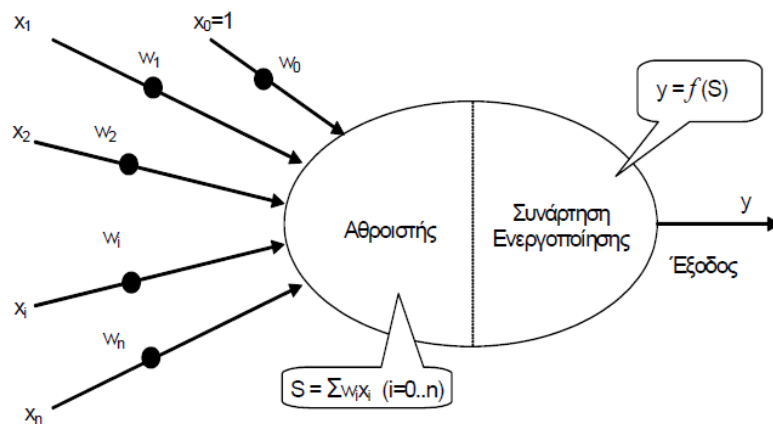


Σχήμα 2.1.1: Διάταξη ενός βιολογικού νευρώνα.

Το νευρωνικό δίκτυο είναι ένα κύκλωμα διασυνδεδεμένων νευρώνων. Στην περίπτωση βιολογικών νευρώνων, πρόκειται για ένα τμήμα νευρικού ιστού. Στην περίπτωση των τεχνητών νευρώνων, πρόκειται για ένα αλγοριθμικό κατασκευάσμα το οποίο ανήκει στο κεφάλαιο της υπολογιστικής νοημοσύνης ή της υπολογιστικής νευροεπιστήμης. Κάθε νευρώνας ενός τεχνητού νευρωνικού δικτύου δέχεται ένα σύνολο αριθμητικών εισόδων, και με βάση αυτές τις εισόδους, των βαρών κάθε σύνδεσης που αντιστοιχεί σε κάθε είσοδο και μια συνάρτηση ενεργοποίησης (σχήμα 2.1.2), παράγει μία έξοδο (σχήμα 2.1.3). Πιο συγκεκριμένα υπολογίζεται το άθροισμα των γινομένων όλων των εισόδων με τα αντίστοιχα βάρη, $S = \sum w_i x_i$ ($i=0..n$), και το άθροισμα αυτό περνά μέσα από τη συνάρτηση ενεργοποίησης, $y=f(S)$, για να μας δώσει την έξοδο του νευρώνα. Αυτή η έξοδος κατευθύνεται σαν είσοδος σε άλλους νευρώνες ή κατευθύνεται προς στο περιβάλλον, σαν έξοδος του νευρωνικού δικτύου. Υπάρχουν τρεις τύποι νευρώνων: οι νευρώνες εισόδου, οι νευρώνες εξόδου και οι κρυφοί νευρώνες.



Σχήμα 2.1.2: Η Λογιστική σιγμοειδής συνάρτηση (αριστερά) και η υπερβολική εφαπτομένη (δεξιά).



Σχήμα 2.1.3: Διάταξη ενός τεχνητού νευρώνα.

Τα τεχνητά νευρωνικά τα χαρακτηρίζει η ικανότητα εκπαίδευσης. Στην προκειμένη περίπτωση η εκπαίδευση έχει να κάνει με την ανανέωση των βαρών των συνδέσεων με απώτερο σκοπό τον εκμηδενισμό του σφάλματος. Υπάρχουν πολλές μέθοδοι εκπαίδευσης, και όλες κατατάσσονται σε δύο κατηγορίες: την επιβλεπόμενη μάθηση (supervised learning) και την μη επιβλεπόμενη μάθηση (unsupervised learning).

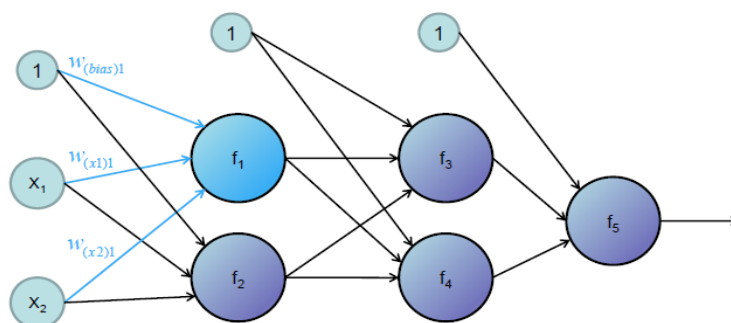
Μέθοδος Ανάστροφης Μετάδοσης Σφάλματος:

Πρόκειται για μια επιβλεπόμενη μέθοδο εκμάθησης, η οποία είναι μια γενίκευση του κανόνα Δέλτα [22], και απαιτεί ένα σύνολο δεδομένων εκπαίδευσης. Οι αλγόριθμοι ανάστροφης μετάδοσης σφάλματος, μεταδίδουν την ποσότητα του σφάλματος με κατεύθυνση από το επίπεδο εξόδου προς το επίπεδο εισόδου. Αρχικά για κάθε νευρώνα εξόδου υπολογίζεται το σφάλμα και γίνεται η ανανέωση των βαρών των εισόδων του. Στη συνέχεια για κάθε

εσωτερικό νευρώνα υπολογίζεται η συμμετοχή του στα σφάλματα των νευρώνων εξόδου και γίνεται παρομοίως η αλλαγή των βαρών του για να μειωθεί το ολικό σφάλμα. Η συμμετοχή ενός νευρώνα στα σφάλματα των νευρώνων του επόμενου επιπέδου του είναι ανάλογη της τρέχουσας εισόδου του και των συντελεστών βαρύτητας που τον συνδέουν με τους νευρώνες του επόμενου επιπέδου. Η διαδικασία αυτή επαναλαμβάνεται μέχρι το ολικό σφάλμα να πάρει μια μικρή αποδεκτή τιμή.

Θα δούμε με την βοήθεια σχημάτων πιο συγκεκριμένα τις δύο βασικές λειτουργίες ενός αλγόριθμου ανάστροφης μετάδοσης σφάλματος εφαρμοσμένη σε ένα πολυεπίπεδο δίκτυο Perceptron (MLM), καθώς και τις συναρτήσεις υπολογισμού σφάλματος και ενημέρωσης βαρών για κάθε νευρώνα του δικτύου. Ένα πολυεπίπεδο δίκτυο Perceptron είναι ένα δίκτυο εμπρόσθιας τροφοδότησης, όπου οι νευρώνες είναι οργανωμένοι σε πολλά επίπεδα, ξεκινώντας από το επίπεδο εισόδου, έχοντας ένα ή περισσότερα κρυφά επίπεδα και καταλήγοντας στο επίπεδο εξόδου. Σημαντικό χαρακτηριστικό αυτής της αρχιτεκτονικής είναι ότι δεν υπάρχουν διασυνδέσεις μεταξύ νευρώνων που ανήκουν στο ίδιο επίπεδο.

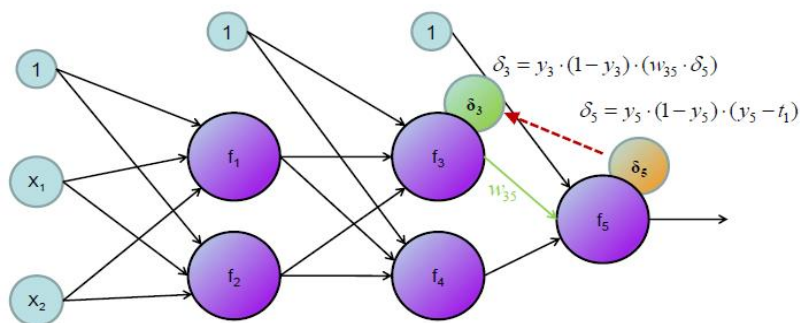
Στο πιο κάτω σχήμα (σχήμα 2.1.4) φαίνεται η πρώτη βασική λειτουργία του αλγόριθμου ανάστροφης μετάδοσης σφάλματος, της φάσης εμπρόσθιας κατεύθυνσης. Στη φάση αυτή, για κάθε νευρώνα υπολογίζεται η έξοδος του, αθροίζοντας όπως βλέπουμε τα γινόμενα όλων των εισόδων που καταλήγουν σε αυτόν με το ανάλογο βάρος της κάθε διασύνδεσης, και περνώντας το αποτέλεσμα αυτό σαν παράμετρο στην συνάρτηση ενεργοποίησης που αναφέραμε πιο πάνω. Η διαδικασία αυτή εφαρμόζεται σε όλους τους νευρώνες, μέχρι να φτάσουμε στους νευρώνες του τελευταίου επιπέδου, των οποίων η έξοδος τους αντιπροσωπεύει και την έξοδο του δικτύου και θα χρησιμοποιηθεί στις επόμενες φάσεις, με σκοπό την μείωση του ολικού σφάλματος.



$$y_1 = f_1(w_{(bias)1} + w_{(x1)1} \cdot x_1 + w_{(x2)1} \cdot x_2)$$

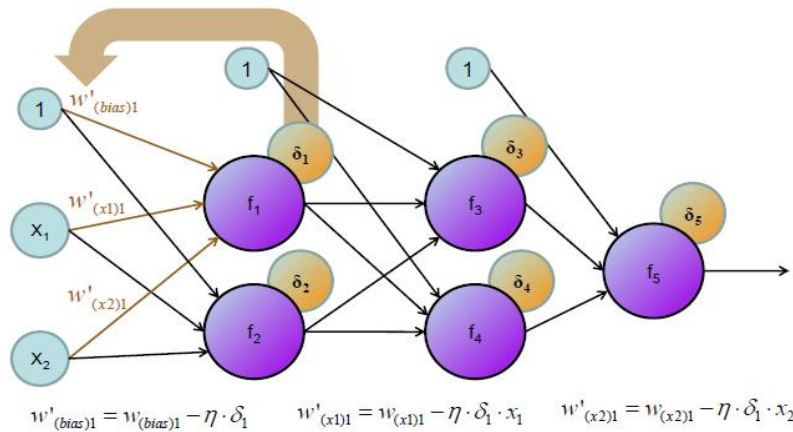
Σχήμα 2.1.4: Φάση εμπρόσθιας κατεύθυνσης

Στη συνέχεια γίνεται ο υπολογισμός του σφάλματος «δ» (δέλτα) για κάθε νευρώνα του δικτύου, που αντιπροσωπεύει την συμμετοχή του κάθε νευρώνα στο ολικό σφάλμα. Η φάση αυτή ονομάζεται ανάστροφη μετάδοση σφάλματος, οπότε ξεκινούμε από το επίπεδο εξόδου και κατευθυνόμαστε προς στο επίπεδο εισόδου. Η εύρεση του σφάλματος για κάθε νευρώνα ενός εσωτερικού επιπέδου παίρνει ως παράμετρο τα σφάλματα όλων των νευρώνων του επόμενου επιπέδου. Επομένως όπως θα δούμε και στο σχήμα (σχήμα 2.1.5) η συνάρτηση εύρεσης του σφάλματος «δ» διαφέρει στους εξωτερικούς νευρώνες από στους εσωτερικούς. Αυτό το σφάλμα «δ» θα χρησιμοποιηθεί στην φάση ενημέρωσης βαρών με σκοπό τη μείωση του ολικού σφάλματος.



Σχήμα 2.1.5. Φάση εύρεσης και ανάστροφης μετάδοσης σφάλματος

Η τελευταία φάση ενός αλγόριθμου ανάστροφης μετάδοσης σφάλματος είναι η ενημέρωση των βαρών των συνδέσεων του δικτύου, ούτως ώστε να μειωθεί το ποσό του ολικού σφάλματος και ουσιαστικά να εκπαιδευτεί το δίκτυο. Η μετατροπή του συντελεστή των βαρών προϋποθέτει ως παραμέτρους την προηγούμενη τιμή βάρους της σύνδεσης, το σφάλμα «δ» του νευρώνα που καταλήγει η κάθε σύνδεση, την τιμή εξόδου του νευρώνα από το προηγούμενο επίπεδο, από τον οποίο ξεκινά η συγκεκριμένη σύνδεση και την τιμή του ρυθμού μάθησης (Learning Rate) η οποία είναι μεταβλητή παράμετρος που δίνουμε στο δίκτυο μας, στην αρχή της εκπαίδευσης. Βλέπουμε πιο κάτω (σχήμα 2.1.6) τις ακριβές συναρτήσεις ενημέρωσης των συντελεστών βαρών.



Σχήμα 2.1.6. Φάση ενημέρωσης βαρών του δικτύου

2.2 Εξελικτικά Συστήματα

Οι εξελικτικοί αλγόριθμοι στην επιστήμη της πληροφορικής ανήκουν στο κεφάλαιο της Τεχνητής Νοημοσύνης, ή πιο συγκεκριμένα της Υπολογιστικής Νοημοσύνης και περιλαμβάνουν προβλήματα συνδυαστικής βελτιστοποίησης. Τα τελευταία χρόνια είναι ένας ευρέως αναπτυσσόμενος τομέας επίλυσης πολλών προβλημάτων. Είναι πιθανοτικές μέθοδοι αναζήτησης, εύκολα στην κατανόηση και την εφαρμογή τους, και μπορούν να εφαρμοστούν και σε δυναμικά συστήματα, στα οποία οι περιορισμοί του προβλήματος μπορούν να μεταβάλλονται με το χρόνο. Ο εξελικτικός υπολογισμός είναι εμπνευσμένος από τους βιολογικούς μηχανισμούς της Δαρβινικής θεωρίας της εξέλιξης. Χρησιμοποιεί μια συνεχή επαναληπτική εξέλιξη, με παρόμοιο τρόπο όπως αυτόν της αύξησης ενός πληθυσμού και στη συνέχεια ο πληθυσμός αυτός επιλέγεται τυχαία μέχρι να φτάσουμε στο επιθυμητό αποτέλεσμα. Οι τεχνικές αυτές εφαρμόζονται όλο και πιο πολύ σε μια ποικιλία από προβλήματα, που κυμαίνονται από πρακτικές εφαρμογές στη βιομηχανία μέχρι την επιστημονική έρευνα.

Μια τέτοια τεχνική επίλυσης προβλημάτων, που ανήκει στον εξελικτικό υπολογισμό, είναι οι Γενετικοί αλγόριθμοι. Οι Γενετικοί αλγόριθμοι υλοποιήθηκαν από τον John Holland [7] το 1975 και γενικά είναι μια μέθοδος αναζήτησης βέλτιστων λύσεων μέσα σε ένα σύνολο υποψηφίων λύσεων, με στόχο την εύρεση των αποδεκτών αποτελεσμάτων. Χαρακτηρίζονται από την απλότητα και την κομψότητα τους ως αλγόριθμοι αναζήτησης, καθώς και από τη δυνατότητα τους να ανακαλύπτουν γρήγορα τις καλές λύσεις δύσκολων προβλημάτων. Είναι

ιδιαίτερα χρήσιμοι σε προβλήματα που έχουν πολλές παραμέτρους και δεν μπορούν να λυθούν με τη χρήση των κλασσικών μεθόδων αναζήτησης. Τα συστατικά ενός γενετικού αλγορίθμου για ένα συγκεκριμένο πρόβλημα είναι:

- Αναπαράσταση λύσεων
- Δημιουργία αρχικού πληθυσμού
- Συνάρτηση καταλληλότητας
- Επιλογή γονέων
- Διαδικασία αναπαραγωγής

Στους Γενετικούς αλγόριθμους οι τιμές για τις παραμέτρους του συστήματος πρέπει να κωδικοποιούνται με τρόπο ώστε να αναπαρασταθούν από μια σειρά χαρακτήρων ή δυαδικών ψηφίων (0/1). Η σειρά αυτή μιμείται το γενετικό κώδικα που υπάρχει στους ζωντανούς οργανισμούς.

Αρχικά, ο αλγόριθμος παράγει πολλαπλά αντίγραφα της σειράς αυτής, συνήθως με τυχαίες τιμές, δημιουργώντας ένα πληθυσμό λύσεων. Κάθε λύση δοκιμάζεται μέσω της συνάρτησης καταλληλότητας, συνάρτηση η οποία μας λέει αριθμητικά πόσο κατάλληλη είναι η κάθε λύση.

Οι καταλληλότερες λύσεις, συγκριτικά με τις υπόλοιπες, αναπαράγονται στην επόμενη γενιά λύσεων και λαμβάνουν μια τυχαία μετάλλαξη. Επαναλαμβάνοντας αυτή τη διαδικασία πολλές φορές, οι τυχαίες μεταλλάξεις σε συνδυασμό με την επιβίωση και αναπαραγωγή των λύσεων που είναι πιο κοντά στο επιθυμητό αποτέλεσμα θα δώσουν τη λύση που θα περιέχει τις τιμές για τις παραμέτρους που ικανοποιούν όσο καλύτερα γίνεται την συνάρτηση ικανότητας.

Κάποια από τα πολλά πλεονεκτήματα των Γενετικών αλγορίθμων είναι ότι δρουν παράλληλα, δηλαδή κινούνται μέσα στο χώρο καταστάσεων με περισσότερα άτομα, έτσι είναι λιγότερο πιθανό να κολλήσουν σε κάποιο τοπικό ακρότατο, σαν κάποιες άλλες μεθόδους. Είναι επίσης, εύκολα υλοποιήσιμοι και προσαρμοστικοί σε διάφορα προβλήματα μεταβάλλοντας μόνο τη συνάρτηση καταλληλότητας και τέλος είναι εύκολα επεκτάσιμοι και εξελίξιμοι. Ένα μειονέκτημα των Γενετικών αλγορίθμων είναι ο υπολογιστικός τους χρόνος.

Είναι πιο αργοί από άλλες μεθόδους, αλλά με τη σημερινή εξέλιξη των υπολογιστών το πρόβλημα αυτό σιγά σιγά εξασθενίζει.

Μερικές εφαρμογές των Γενετικών αλγορίθμων είναι:

- η καλύτερη δυνατή οργάνωση ενός σχολικού ωραρίου
- η μελέτη της βέλτιστης κατανομής ενός δικτύου από πλατφόρμες πετρελαίου
- εξερεύνηση των δυναμικών βιολογικών διαδικασιών, και της θεωρίας της εξέλιξης.

2.3 Συστήματα Ασαφούς Λογικής

Η ασαφής λογική (fuzzy logic) είναι μια επέκταση της κλασσικής αριστοτέλειας λογικής. Μια πρόταση μπορεί να είναι αληθής με κάποιο βαθμό αληθείας, και όχι απλά αληθής ή ψευδής. Υπάρχει δηλαδή μια έλλειψη ακρίβειας που ερμηνεύεται ως ασάφεια. Η ιδέα αυτή απετέλεσε επανάσταση στη θεωρία της λογικής, γιατί ξέφυγε από το μοντέλο που κυριαρχούσε για χρόνια, δηλαδή το μοντέλο του «αληθές-ψευδές». Η έννοια της ασαφούς λογικής προσδιορίστηκε από τον Lotfi A. Zadeh[21], καθηγητής στο Πανεπιστήμιο της Καλιφόρνιας στο Berkeley, περίπου το 1960 και παρουσιάστηκε ως τρόπος επεξεργασίας δεδομένων που δεν ήταν ακριβή. Έτσι μπαίνουμε σε μια καινούργια λογική που μας λέει ότι υπάρχουν προτάσεις που θα ισχύουν με μια πιθανότητα. Η ασαφής λογική προσπαθεί να προσομοιώσει τις ιδιότητες του ανθρώπινου μυαλού και να παίρνει αποφάσεις από ένα ασαφή σύνολο δεδομένων μέσα σε ένα περιβάλλον αβεβαιότητας και ανακρίβειας. Η ανάγκη αυτή γεννήθηκε από το γεγονός ότι οι προσεγγίσεις που βασίζονται στην κανονική λογική δεν μπορούσαν να αναπαραστήσουν τη γνώση με επιτυχία, κυρίως σε δύσκολα συστήματα με ασαφείς έννοιες. Η ασαφής λογική έχει πάρα πολλές εφαρμογές σε διάφορες επιστήμες όπως πληροφορική, φυσική και ψυχολογία.

Η ασαφής λογική έχει πολλά πλεονεκτήματα, γιατί παρόλο που λειτουργούν σε περιβάλλον ασάφειας δίνουν αποτελέσματα που έχουν νόημα και είναι ευκατανόητα από την άνθρωπο. Είναι δηλαδή ένα εργαλείο πολύ κοντά στην ανθρώπινη λογική και στον ανθρώπινο τρόπο σκέψης και χρησιμοποιείται για την λήψη αποφάσεων. Χρησιμοποιείται σε πολύπλοκα συστήματα στα οποία είναι δύσκολη ή αδύνατη η εύρεση κάποιου μαθηματικού μοντέλου για αυτά. Άλλο σημαντικό πλεονέκτημα της είναι η μεγάλη ανεκτικότητα που έχει σε ανακριβή δεδομένα.

Η ασαφής λογική έχει καταφέρει να λύσει προβλήματα για τα οποία μαθηματικά μοντέλα αποτυγχάνουν, αλλά η εύρεση των κανόνων και συναρτήσεων συμμετοχής που απαιτούνται για την αποτελεσματική περιγραφή των συστημάτων απαιτεί σημαντική έρευνα και προσπάθεια ώστε να έχουμε καλά αποτελέσματα. Επίσης, η ασαφής λογική χρησιμοποιεί ευριστικούς αλγόριθμους, των οποίων οι λύσεις που δίνουν δεν μπορούν να ανταποκριθούν πάντα αποτελεσματικά. Τέλος, η ασαφής λογική δεν έχει ιδιαίτερη ικανότητα γενίκευσης όπως τα νευρωνικά δίκτυα, ιδιότητα που είναι πολύ σημαντική, ώστε να μπορεί το μοντέλο να ανταποκριθεί σε καινούριες καταστάσεις και άγνωστα δεδομένα. Ο συνδυασμός της με τα νευρωνικά δίκτυα και τους εξελικτικούς αλγορίθμους οδήγησε στη δημιουργία συστημάτων με ικανότητες μάθησης, γενίκευσης και συμπερασμού.

Τα Ασαφή Συστήματα στηρίζονται στη θεωρία της ασαφούς λογικής και των ασαφών συνόλων. Είναι συστήματα που λειτουργούν όπως το ανθρώπινο μυαλό, σχετίζονται με την έννοια της μερικής αλήθειας και αντικαθιστούν τις τιμές αληθείας με βαθμούς αληθείας. Βρίσκουν εφαρμογή κυρίως σε θέματα αυτομάτου ελέγχου στη βιομηχανική παραγωγή, όμως η θεωρία τους έχει διαδοθεί και σε άλλες περιοχές, όπως στην παραγωγή φωτογραφικών μηχανών, βιντεοκαμερών, πλυντηρίων, κλιματιστικών, στην πραγματοποίηση δημοσκοπήσεων, στην οργάνωση επιχειρήσεων, στην ψυχολογία, την ιατρική, αλλά και τη σεισμολογία.

Κεφάλαιο 3

Προηγούμενες Μελέτες

3.1 Διπλωματικές εργασίες στο Πανεπιστήμιο Κύπρου

3.2 Εργασίες στο Πανεπιστήμιο Κύπρου

3.3 Εργασίες σε διεθνές επίπεδο

3.1 Διπλωματικές εργασίες στο Πανεπιστήμιο Κύπρου

Έχουν γίνει αρκετές διπλωματικές εργασίες στο Πανεπιστήμιο μας σχετικά με την πρόβλεψη τρισωμιών χρησιμοποιώντας συστήματα υπολογιστικής νοημοσύνης. Αυτή η εργασία [24] είχε σκοπό να βρει ένα νευρωνικό δίκτυο που θα κάνει σωστή πρόβλεψη αυτών των προβλημάτων. Χρησιμοποίησε 34,528 περιπτώσεις περιστατικών, τα οποία επεξεργάστηκαν, έγινε μελέτη διαφόρων αρχιτεκτονικών μέχρι να βρεθεί αυτή με τα καλύτερα δυνατά αποτελέσματα. Μελετήθηκαν οι τρισωμίες 13, 18, 21 η τριπλοειδής και το σύνδρομο Turner. Η προσπάθεια αυτή κατέληξε σε μια αρχιτεκτονική νευρωνικού δικτύου που έφερε ικανοποιητικά αποτελέσματα, γύρω στο 76% στα δεδομένα εκπαίδευσης και 72% στα δεδομένα επαλήθευσης. Τα ποσοστά επιτυχίας ήταν πολύ χαμηλά έτσι το σύστημα δεν θεωρείται ιδανικό για πραγματική χρήση, γιατί σε τέτοιου είδους περιπτώσεις, όπου κρίνονται ζωές, τα ποσοστά επιτυχίας των προβλέψεων πρέπει να αγγίζουν το 100%.

Ακόμα μια εργασία [23] έχει γίνει με αυτόν τον σκοπό. Αυτή η εργασία χρησιμοποίησε μια βάση δεδομένων με συνολικά 51,208 περιστατικά εγκύων γυναικών. Στη συνέχεια

φτιάχθηκαν 231 μοντέλα με διαφορετική δομή, με σκοπό την εξαγωγή συμπερασμάτων για την συνεισφορά και την ευαισθησία των χαρακτηριστικών εισόδου. Τα αποτελέσματα έδειξαν ότι τη μεγαλύτερη συνεισφορά είχαν η αυχενική διαφάνεια του εμβρύου και η ροή φλεβώδους πόρου, ενώ λιγότερη αλλά σημαντική συνεισφορά είχαν η πρωτεΐνη A και η ηλικία της μητέρας.

3.2 Εργασίες στο Πανεπιστήμιο Κύπρου

Μια μελέτη [15] έχει γίνει στο Πανεπιστήμιο Κύπρου για την σημασία που έχουν στην πρόβλεψη των χρωμοσωμικών ανωμαλιών η Χοριακή Γονιαδοτροπίνη (b-hCG) και η Πρωτεΐνη A (PAPP-A) χρησιμοποιώντας διάφορες δομές Νευρωνικών Δικτύων. Τα δεδομένα της βάσης δεδομένων που χρησιμοποιήθηκε ήταν ιδιαίτερα ανισόροπα. Περιείχε 35687 περιπτώσεις εγκύων γυναικών εκ των οποίων οι 35058 ήταν απολύτως φυσιολογικές χωρίς οποιαδήποτε χρωμοσωμική ανωμαλία, ενώ στις υπόλοιπες 629 υπήρχε κάποιο είδος χρωμοσωμικής ανωμαλίας. Το δίκτυο εκπαιδεύτηκε με αυτά τα δεδομένα και στη συνέχεια δοκιμάστηκαν 8181 άγνωστες περιπτώσεις για την επαλήθευση του. Στο σύνολο αυτό υπήρχαν 76 προβληματικές περιπτώσεις. Το σύστημα εκπαιδεύεται με τη χρήση 8 χαρακτηριστικών στην είσοδο, τα οποία που θεωρούνται ότι είναι τα πιο σημαντικά στον προσδιορισμό του κινδύνου εμφάνισης αυτών των ανωμαλιών. Στη συνέχεια, η Χοριακή Γονιαδοτροπίνη (b-hCG) και η Πρωτεΐνη A (PAPP-A) απομακρύνθηκαν από τις εισόδους, προκειμένου να διαπιστωθεί η επίδραση τους στο αποτέλεσμα. Τα καλύτερα αποτελέσματα λήφθηκαν όταν χρησιμοποιήθηκε ένα πολυστρωματικό νευρωνικό δίκτυο με δύο κρυμμένα στρώματα και διαπιστώθηκε ότι τόσο η Χοριακή Γονιαδοτροπίνη όσο και η Πρωτεΐνη A είναι απαραίτητα χαρακτηριστικά προκειμένου να επιτευχθεί υψηλό ποσοστό επιτυχίας.

Μια άλλη μελέτη που έγινε στο Πανεπιστήμιο [14] ασχολήθηκε περεταίρω με την συμβολή και την σημασία που έχουν περισσότερα χαρακτηριστικά στην μη επεμβατική πρόβλεψη χρωμοσωμικών ανωμαλιών, χρησιμοποιώντας κάποια μοντέλα τεχνητών νευρωνικών δικτύων. Οι παράμετροι που είχαν ερευνηθεί σε αυτή την εργασία είναι το προηγούμενο ιατρικό ιστορικό της εγκύου μητέρας, το ρινικό οστό, η τριγλώχινα ροή, η ροή φλεβώδους πόρου, η πρωτεΐνη A, η χοριακή γονιαδοτροπίνη, το κεφαλουραίο μήκος, οι αλλαγές στην αυχενική διαφάνεια και η ηλικία της μητέρας. Η μελέτη έφτασε στα συμπεράσματα ότι η αυχενική διαφάνεια είναι ο πιο σημαντικός παράγοντας για την συνολική πρόβλεψη, ενώ το

κεφαλουραίο μήκος το λιγότερο σημαντικό. Συμπεράθηκε ακόμα ότι το προηγούμενο ιατρικό ιστορικό της εγκύου μητέρας δεν είναι σημαντικός παράγοντας για την πρόβλεψη προβληματικών περιπτώσεων αλλά ότι το ρινικό οστό, η τριγλώχινια ροή και ο φλεβώδης πόρος συμβάλλουν σημαντικά στην πρόβλεψη της τρισωμίας 21, αλλά όχι στην πρόβλεψη των φυσιολογικών περιπτώσεων. Η πρωτεΐνη A, η χοριακή γονιαδοτροπίνη και η ηλικία της μητέρας είναι ενδιάμεσης σημασίας. Επίσης, έγινε και μία ανάλυση ευαισθησίας των παραμέτρων πρωτεΐνη A, χοριακή γονιαδοτροπίνη, κεφαλουραίο μήκος, αυχενική διαφάνεια και ηλικία της μητέρας. Τα αποτελέσματα έδειξαν ότι το κεφαλουραίο μήκος και η αυχενική διαφάνεια ευαισθητοποιούνται περεταίρω όταν οι τιμές τους μειώνονται, η πρωτεΐνη A ευαισθητοποιείται όταν η τιμή της αυξάνεται, ενώ η χοριακή γονιαδοτροπίνη μένει ανεπηρέαστη από τις διακυμάνσεις της τιμής της.

Ακόμα μια εργασία που έχει γίνει από το Πανεπιστήμιο Κύπρου [16] είχε εστιάσει στην πρόγνωση της τρισωμίας 21 κατά την πρώτη τριμηνία, χρησιμοποιώντας επίσης τεχνητά νευρωνικά δίκτυα. Η τρισωμία 21 ή σύνδρομο Down, όπως ανέφερα και πριν, προκαλείται από παρουσία ενός επιπλέον 21^{ου} χρωμοσώματος. Στην εργασία αυτή χρησιμοποιήθηκαν δεδομένα από εξετάσεις υπέρηχων στα έμβρυα σε έξυπνα συστήματα με στόχο την πρόβλεψη της πάθησης. Έχουν γίνει δοκιμές με πολλά νευρωνικά δίκτυα εμπρόσθιας τροφοδότησης μέχρι να βρεθεί η καλύτερη συνταγή με την πιο επιτυχημένη πρόβλεψη. Η βάση δεδομένων περιείχε 23.513 περιπτώσεις και η κάθε έγκυος είχε 19 παραμέτρους, από τις οποίες οι 11 θεωρήθηκαν οι πιο καθοριστικές στον υπολογισμό του κινδύνου γενετικής ανωμαλίας. Τα αποτελέσματα της προσπάθειας αυτής ήταν πολύ ψηλά, έχοντας ποσοστά πάνω από 90% και στα δεδομένα εκπαίδευσης, αλλά και επαλήθευσης.

3.3 Εργασίες σε διεθνές επίπεδο

Μελέτες έχουν γίνει και στο εξωτερικό με σκοπό την εύρεση αξιόπιστου μη επεμβατικού τρόπου πρόγνωσης γενετικών προβλημάτων. Ξεκινώντας την αναφορά μου σε αυτές, μία μελέτη [8] που έγινε, είχε σκοπό την διάγνωση της Τρισωμίας 21 κατά το πρώτο τρίμηνο σε δύο στάδια, με εκτίμηση υπερηχογραφήματος και βιοχημικών ελέγχων. Στο πρώτο στάδιο της εργασίας γίνεται υπολογισμός του κινδύνου με βάση τον συνδυασμό χαρακτηριστικών όπως η ηλικία της μητέρας, η αυχενική διαφάνεια του εμβρύου, η παρουσία ή η απουσία του ρινικού οστού, η ροή του φλεβώδους πόρου, η τριγλώχινια ροή. Στη συνέχεια, βιοχημικός

έλεγχος γινόταν μόνο στις περιπτώσεις που ανιχνευόταν τουλάχιστο ενδιάμεσος κίνδυνος κατά το πρώτο στάδιο. Μελετήθηκε ακόμη και η απόδοση της πολιτικής η οποία προέβλεπε σαν πρώτο στάδιο την εκτίμηση του κινδύνου σύμφωνα με την ηλικία της μητέρας και βιοχημικές εξετάσεις. Σε δεύτερο στάδιο, δεδομένης κάποιας πιθανότητας κινδύνου, πραγματοποιούνταν υπερηχογραφικές εξετάσεις. Τα δεδομένα αποτελούνταν από 19,614 περιστατικά φυσιολογικών εγκυμοσυνών και 122 περιπτώσεις τρισωμίας 21. Η καλύτερη επίδοση στις προγνώσεις των γενετικών ανωμαλιών επιτεύχθηκε ακολουθώντας πρώτα την πολιτική όπου εξεταζόταν το μέγεθος του κινδύνου σύμφωνα με την ηλικία της μητέρας, την αυχενική διαφάνεια του εμβρύου και είτε της τριγλώχινας ροής είτε της ροής φλεβόδους πόρου. Σε περιπτώσεις που ο κίνδυνος ήταν τουλάχιστο ενδιάμεσος, πιο συγκεκριμένα από 1 προς 51 μέχρι 1 προς 1000, ακολούθησε βιοχημική εξέταση. Στις δοκιμές όπου το πρώτο στάδιο προέβλεπε βιοχημικές εξετάσεις και οι υπερηχογραφικές εξετάσεις λαμβάνονταν υπόψη στο δεύτερο στάδιο, δεν είχαν τα ίδια ποσοστά επιτυχίας. Έτσι, η μελέτη κατέληξε ότι μια αποτελεσματική πρόβλεψη της τρισωμίας 21 κατά την πρώτη τριμηνία επιτυγχάνεται με υπερηχογραφικές εξετάσεις κατά το πρώτο στάδιο, από τις οποίες επιλέγεται το 20% των περιπτώσεων, ως περιπτώσεις με ενδιάμεσο κίνδυνο, και πάνω σε αυτές εφαρμόζονται βιοχημικές εξετάσεις.

Ακόμα μια εργασία [18] που διεκπεραιώθηκε με σκοπό την πρόβλεψη γενετικών ανωμαλιών πρόωρα, είχε ως σκοπό την ανίχνευση της τρισωμίας 21, χρησιμοποιώντας 200,868 συνολικά κυήσεις, εκ των οποίων οι 871 ήταν προβληματικές. Η εργασία αυτή έχει δείξει ότι μια αυξημένη αυχενική διαφάνεια του εμβρύου μπορεί να προσδιορίσει το 76.8% αυτών με τρισωμία 21. Όταν χρησιμοποιήθηκε η αυχενική διαφάνεια σε συνδυασμό με τη χοριακή γονιαδοτροπίνη και την πρωτεΐνη Α σε δεδομένα 44,613 περιπτώσεων, εκ των οποίων 215 είχαν προσβληθεί από την τρισωμία 21, είχαμε ένα ποσοστό ανίχνευσης του 87%. Μελέτες από εξειδικευμένα κέντρα, με 15,822 περιπτώσεις, συμπεριλαμβανομένων και 397 με τρισωμία 21, έχουν δείξει ότι η απουσία του ρινικού οστού μπορεί να προσδιορίσει το 69% των εμβρύων με ανωμαλία στο χρωμόσωμα 21. Έχει εκτιμηθεί ακόμα ότι ο συνδυασμός υπερηχογραφικών εξετάσεων και των αποτελεσμάτων μιας εξέτασης που αναλύει κάποιες μητρικές ορμόνες [13], μπορεί να αναγνωρίσει το 97% των εμβρύων με τρισωμία 21.

Αυτή η εργασία [12] θέλησε να διερευνήσει τη σημασία που έχει η ροή φλεβόδους πόρου σε συνδυασμό με την ηλικία της μητέρας, την αυχενική διαφάνεια του εμβρύου, τον καρδιακό παλμό του εμβρύου, την χοριακή γονιαδοτροπίνη και την πρωτεΐνη Α. Τα δεδομένα περιλάμβαναν 19.614 φυσιολογικά έμβρυα, 122 με τρισωμία 21, 36 με τρισωμία 18, 20 με

τρισωμία 13 και 8 με σύνδρομο Turner. Τα αποτελέσματα χωρίς να ληφθεί υπόψη η ροή φλεβώδους πόρου είχε ποσοστά επιτυχίας γύρω στο 65%, ενώ εντάσσοντας το δεδομένο αυτό στην προσπάθεια διάγνωσης έφερε ποσοστά ανίχνευσης 96%, 92%, 100% και 100% για τις τρισωμίες 21, 18, 13 και συνδρόμου Turner, αντίστοιχα. Έτσι η μελέτη κατέληξε στο συμπέρασμα ότι η εκτίμηση της ροής φλεβώδους πόρου βελτιώνει την απόδοση επιτυχούς πρόβλεψης ανευπλοειδιών.

Πάρα πολλές μελέτες έχουν γίνει, με την μέθοδο της μελέτης αλληλούχισης του μητρικού DNA με σκοπό την εύρεση δειγμάτων του DNA του εμβρύου. Αυτή η εργασία [4] στηρίχθηκε σε αυτή τη μέθοδο και προσπάθησε να ανιχνεύσει τυχόν τρισωμία στο μωρό. Χρησιμοποιήθηκε ένα σύνολο από 480 δείγματα από γυναίκες με υψηλό κίνδυνο. Τα αποτελέσματα της εργασίας είχαν ικανοποιητικά αποτελέσματα, γεγονός που επαληθεύει την υπόθεση ότι η μελέτη αλληλουχίας του DNA θα μπορούσε να γίνει μια αξιόπιστη μέθοδος για μη επεμβατική ανίχνευση εμβρύου με τρισωμία 21.

Σε αυτή την εργασία [2] έγινε προσπάθεια για προγεννητική εξέταση με επιλεκτική ανάλυση ελεύθερων κυττάρων DNA στο μητρικό αίμα για την αξιολόγηση του εμβρύου με τρισωμία 21 και 18. Χρησιμοποιήθηκαν 298 εγκυμοσύνες εκ των οποίων οι 39 είχαν πρόβλημα τρισωμίας 21 και οι 7 είχαν πρόβλημα τρισωμίας 18. Αυτά τα δεδομένα αναλύθηκαν χρησιμοποιώντας μια σύγχρονη μέθοδο, που ονομάζεται ψηφιακή ανάλυση επιλεγμένων περιοχών[17]. Η προσπάθειες έγιναν σε μεγάλα βάθη αλληλουχιών, φτάνοντας τις 620,000 ανά δείγμα. Τα αποτελέσματα της εργασίας ήταν θετικά βγάζοντας το συμπέρασμα ότι η ψηφιακή ανάλυση επιλεγμένων περιοχών δίνει εξαιρετικά ακριβής με επεμβατικές αξιολογήσεις του εμβρύου για τυχόν χρωμοσωμικές διαταραχές με χαμηλό οικονομικό κόστος.

Προσπάθειες για πρόβλεψη εμβρυϊκών ανωμαλιών έχουν γίνει και με στατιστικές μεθόδους όπως συμβαίνει σε αυτή τη μελέτη [11], όπου γίνεται η προσπάθεια με τη χρήση πλάσματος της πρωτεΐνης A κατά το πρώτο τρίμηνο. Χρησιμοποιήθηκαν 475 περιπτώσεις και έγινε σύγκριση μεταξύ δύο διαφορετικών προσεγγίσεων για στατιστική ανάλυση, της Γκαουσιανής κατανομής και της λογιστικής παλινδρόμησης. Και οι δύο μέθοδοι είχαν καλά αποτελέσματα ανάλογα με τη φύση των δεδομένων. Τα ποσοστά ανίχνευσης κυμαίνονταν από 42% μέχρι 70%.

Κεφάλαιο 4

Ανάλυση θέματος και εξήγηση

4.1 Ορισμός προβλήματος

4.2 Ερευνητική ερώτηση

4.3 Ανάλυση χαρακτηριστικών

4.4 Παραδοσιακές μέθοδοι αντιμετώπισης

4.5 Σύγκριση και συζήτηση μεθόδων

4.1 Ορισμός Προβλήματος

Το ερευνητικό πρόβλημα με το οποίο ασχολήθηκα είναι η πρόβλεψη γενετικών ανωμαλιών, χρησιμοποιώντας έξυπνα συστήματα υπολογιστικής νοημοσύνης. Πιο συγκεκριμένα δημιουργήσα μοντέλα νευρωνικών δικτύων που ξεχωρίζουν τα έμβρυα με χρωμοσωμικές διαταραχές από αυτά που δεν έχουν κάποια χρωμοσωμική διαταραχή. Για να προβλεφθούν ή να διαγνωστούν χρωμοσωμικές διαταραχές εντός της πρώτης τριμηνίας της εγκυμοσύνης υπάρχουν αρκετοί τρόποι, οι οποίοι χωρίζονται σε επεμβατικούς και μη επεμβατικούς. Η κάθε κατηγορία μεθόδου διάγνωσης έχει κάποια πλεονεκτήματα και κάποια μειονεκτήματα.

Στην εργασία μου προσέγγισα το πρόβλημα αυτό με την χρήση έξυπνων συστημάτων. Χρησιμοποίησα μια βάση δεδομένων με μεγάλο αριθμό διαφορετικών περιπτώσεων, η οποία παραχωρήθηκε από το Fetal Medicine Foundation [5]. Σε αυτή τη βάση δεδομένων υπάρχουν περιπτώσεις από πέντε διαφορετικές χρωμοσωμικές ανωμαλίες και μία όπου υπάρχουν περιπτώσεις φυσιολογικών εμβρύων. Στην εργασία μου εξέτασα μόνο τις περιπτώσεις για τρισωμία 21.

Χρησιμοποιώντας την βάση δεδομένων, δημιούργησα μοντέλα με Νευρωνικά Δίκτυα και τα εκπάιδευσά με διάφορες παραμέτρους όπως τον αριθμό των εποχών εκπαίδευσης, την συνάρτηση ενεργοποίησης των νευρώνων, το ρυθμό μάθησης, την ορμή και με διάφορα χαρακτηριστικά, με σκοπό να φτάσω σε επιθυμητά ποσοστά επιτυχίας στην πρόβλεψη χρωμοσωμικών διαταραχών, βάζοντας σαν είσοδο στο σύστημα κάποια χαρακτηριστικά από ιατρικές εξετάσεις της μητέρας και του εμβρύου. Αυτή είναι μια μη επεμβατική μέθοδος με την χρήση έξυπνων συστημάτων όπου τα πλεονεκτήματα της είναι το χαμηλό κόστος και η ταχύτητα με την οποία δίνει αποτελέσματα. Η πρόκληση για την ανάπτυξη αυτής της μεθόδου είναι να έχει το ίδιο αξιόπιστα αποτελέσματα με τις επεμβατικές μεθόδους, ούτως ώστε να χρησιμοποιείται ευρέως.

Χρωμοσωμικές ανωμαλίες:

Το χρωμόσωμα είναι μια οργανωμένη δομή DNA και πρωτεϊνών που βρίσκεται στα κύτταρα. Κάθε κύτταρο του ανθρώπινου οργανισμού υπό φυσιολογικές συνθήκες περιέχει 23 ζεύγη, 46 δηλαδή χρωμοσώματα, εκ των οποίων τα 23 προέρχονται από τον πατέρα και 23 από την μητέρα. Τα πρώτα 22 ζευγάρια λέγονται αυτοσωμικά και είναι ίδια στους άνδρες και γυναίκες. Το 23^ο ζευγάρι χρωμοσωμάτων αποτελείται από τα λεγόμενα φυλετικά χρωμοσώματα, και καθορίζουν το φύλο του ατόμου. Οι χρωμοσωμικές ανωμαλίες είναι αλλαγές στη δομή ενός ή περισσότερων χρωμοσωμάτων. Οι πιο βασικές μεταλλάξεις χρωμοσωμάτων είναι οι εξής:

- Διαγραφή: απώλεια μέρους ενός χρωμοσώματος
- Διπλασιασμός: επιπλέον αντίγραφα ενός μέρους ενός χρωμοσώματος
- Αντιστροφή: αντιστροφή κατεύθυνσης ενός μέρους ενός χρωμοσώματος
- Μετατόπιση: μέρος ενός χρωμοσώματος αποσπάται και επισυνάπτει ένα άλλο χρωμόσωμα. Αυτή η μετάλλαξη έχει ως αποτέλεσμα την δημιουργία μονοσωμιών ή τρισωμιών.

Πιο συγκεκριμένα σε ένα έμβρυο με τρισωμία παρουσιάζεται ένα χρωμόσωμα 3 φορές αντί 2 που είναι το φυσιολογικό. Η πιο γνωστή χρωμοσωμική ανωμαλία είναι το σύνδρομο «Down». Το σύνδρομο αυτό οφείλεται στην παρουσία ενός επιπλέον χρωμοσώματος στη θέση 21, εξού και η εναλλακτική του ονομασία ως τρισωμία 21. Αντίστοιχα το σύνδρομο «Patau» ονομάζεται και

τρισωμία 13 λόγω της επιπλέον ύπαρξης του χρωμοσώματος 13 και το σύνδρομο «Edward» που παρομοίως ονομάζεται και τρισωμία 18.

Σύνδρομο «Patau» (Τρισωμία 13)

Το σύνδρομο Patau, γνωστό και ως Τρισωμία 13, είναι μια σχετικά σπάνια χρωμοσωμική ανωμαλία. Κατά την ανωμαλία αυτή το άτομο γεννιέται με ένα επιπλέον χρωμόσωμα στο 13^ο ζεύγος. Είναι η τρίτη πιο συνηθισμένη τρισωμία μετά την Τρισωμία 21 και την Τρισωμία 18. Η συχνότητα με την οποία εμφανίζεται το σύνδρομο Patau φτάνει περίπου την 1 ανά 5.000 με 10.000 γεννήσεις (1:5.000-10.000). Στις πλείστες περιπτώσεις το έμβρυο δεν επιβιώνει και αποβάλλεται. Φαίνεται να αυξάνεται ο κίνδυνος εκδήλωσης του συνδρόμου με την αύξηση της ηλικίας της μητέρας ειδικά αν πρόκειται για γυναίκα άνω των 30 ετών. Το σύνδρομο Patau προκαλεί βαριά νοητική καθυστέρηση και σοβαρές μαθησιακές δυσκολίες. Υπάρχουν ακόμα πολλά συμπτώματα που σχετίζονται με την υγεία του παιδιού όπως η μικροφθαλμία, το μικρό κεφάλι, κρανίο χωρίς δέρμα, μικρό βάρος κατά τη γέννηση, κώφωση και καρδιακές ανεπάρκειες. Όπως το σύνδρομο Edward, έτσι και το Patau συνήθως επιφέρει το θάνατο των εμβρύων τις περισσότερες φορές προτού γεννηθούν ενώ όσα καταφέρνουν να γεννηθούν ζουν για λίγες μόνο μέρες ή μέχρι έναν μήνα. Μόνο ένα μικρό ποσοστό, κάτω του 10%, ζουν το πολύ για έναν χρόνο. Δεν υπάρχει θεραπεία για τη σπάνια αυτή πάθηση.

Σύνδρομο «Edward» (Τρισωμία 18)

Το σύνδρομο Edwards, ή αλλιώς Τρισωμία 18, είναι επίσης ένα σπάνιο σύνδρομο που πιθανόν να καταλήξει σε σοβαρή αναπηρία και αποτελεί αιτία θανάτου των βρεφών που φέρουν το σύνδρομο αυτό. Στην Τρισωμία 18 υπάρχει ένα επιπλέον χρωμόσωμα στο 18^ο ζεύγος, επηρεάζει όλα τα συστήματα του σώματος και προκαλεί συγκεκριμένα χαρακτηριστικά προσώπου. Η συχνότητα εμφάνισης του είναι περίπου 1 ανά 6.000-8.000 γεννήσεις. Γύρω στο 95% των εμβρύων πεθαίνουν πριν τη γέννηση. Η Τρισωμία 18 επηρεάζει σε μεγάλο βαθμό όλα τα οργανικά συστήματα του σώματος, τον εγκέφαλο, το πρόσωπο, την καρδιά και τα οστά. Τους πρώτους μήνες της ζωής τους, βρέφη που φέρουν το σύνδρομο χρειάζονται ειδική φροντίδα από ιατρικό προσωπικό. Κατά τη γέννηση το σύνδρομο γίνεται αντιληπτό από την εξωτερική εμφάνιση του νεογέννητου, αλλά μπορεί να διαγνωστεί και προγεννητικά από εξετάσεις με υπερήχους ή με αμνιοκέντηση. Λόγω της περιπλοκότητας των συμπτωμάτων που συνοδεύουν το σύνδρομο Edwards, με κύριο σύμπτωμα την καρδιακή ανεπάρκεια και τον κίνδυνο λοιμώξεων, περίπου το 5% των νηπίων επιζεί κατά τον πρώτο χρόνο ζωής.

Σύνδρομο «Down» (Τρισωμία 21)

Το σύνδρομο Down, αλλιώς Τρισωμία 21 ή Τρισωμία G, είναι μια χρωμοσωμική ανωμαλία, που προκαλείται όταν το άτομο γεννιέται με ένα επιπλέον χρωμόσωμα στο 21^ο ζεύγος. Το σύνδρομο αποτελείται από ένα σύνολο χαρακτηριστικών, τα οποία υπάρχουν εκ γενετής στους φορείς του, και αφορούν παρεκκλίσεις στη σωματική διάπλαση, τη νοητική ανάπτυξη και την ψυχοκοινωνική εξέλιξή τους. Η συχνότητα του συνδρόμου υπολογίζεται να είναι γύρω στις μία ανά 700 γεννήσεις και είναι η συχνότερη γενετική αιτία νοητικής υστέρησης. Σημαντικός παράγοντας που επηρεάζει την αναλογία αυτή είναι η ηλικία της μητέρας, εφόσον στην ηλικία των 20 η συχνότητα κυμαίνεται περίπου στο 1 προς 2000 ενώ μετά τα 40 ξεπερνά το 1 προς 100.

Προγεννητικά είναι δυνατή η διάγνωση πιθανής ύπαρξης του συνδρόμου Down στις έγκυες γυναίκες με επεμβατικές μεθόδους, από τα πρώτα κιόλας στάδια της κύησης, αλλά η διάγνωση μπορεί να γίνει και μετά τη γέννηση. Υπάρχουν ακόμα και κάποια βασικά χαρακτηριστικά που μπορεί να παρατηρηθούν στα παιδιά και να συμπεραθεί ότι είναι φορείς του συνδρόμου. Δεν υπάρχει ολική θεραπεία του συνδρόμου, αλλά γίνονται παρεμβάσεις σε όποιες από τις εκδηλώσεις του συνδρόμου είναι εφικτό. Το ανοσοποιητικό σύστημα των φορέων είναι ιδιαίτερα επιρρεπές σε ασθένειες και σπάνια ξεπερνούν τα 50 χρόνια ζωής.

Σύνδρομο «Triploid»

Το σύνδρομο Triploid είναι μια επίσης πολύ σπάνια χρωμοσωμική διαταραχή. Κατά το σύνδρομο αυτό το έμβρυο έχει 3 χρωμοσώματα αντί για 2 σε όλα τα 'ζεύγη', έχοντας στο σύνολο 69 χρωμοσώματα αντί για 46. Συνήθως τα έμβρυα με το σύνδρομο Triploid αποβάλλονται πρόωρα, αλλιώς αν γεννηθούν επιβιώνουν για διάστημα περίπου 5 μηνών. Αυτά τα παιδιά αντιμετωπίζουν κάποια κοινά συμπτώματα όπως χαμηλό ανάστημα, πρόβλημα όρασης, ευρύς θώρακας και πεταλοειδές νεφρός.

Σύνδρομο «Turner»

Το σύνδρομο Turner είναι μια γενετική ανωμαλία που προκαλείται από την έλλειψη ενός φυλετικού χρωμοσώματος σε θηλυκά άτομα και πρόκειται για τη συχνότερη μονοσωμία που συναντούμε σε αποβαλλόμενα έμβρυα. Γεννήσεις κοριτσιών με σύνδρομο Turner εμφανίζονται με συχνότητα περίπου στο 1 προς 2500. Το σύνδρομο μπορεί να διαγνωστεί

προγεννητικά με αμνιοπαρακέντηση και άλλες μεθόδους. Συνήθως τα χαρακτηριστικά του συνδρόμου είναι εμφανή από τη γέννηση του κοριτσιού, αλλά γίνονται εντονότερα κατά την εφηβεία. Το χρωμόσωμα που λείπει δεν μπορεί να αντικατασταθεί και να θεραπευτεί το σύνδρομο, αλλά με την χορήγηση κάποιων ορμονών οι φορείς αυξάνουν το ύψος και τα θηλυκά χαρακτηριστικά τους.

4.2 Ερευνητική ερώτηση

- Θα μπορούσαν έξυπνα πληροφοριακά συστήματα να εκπαιδευτούν ούτως ώστε να είναι σε θέση να διαγνώσουν έγκαιρα διάφορες γενετικές ανωμαλίες του εμβρύου, χρησιμοποιώντας μόνο δεδομένα από μη επεμβατικές εξετάσεις και χαρακτηριστικά της μητέρας;
- Ποια χαρακτηριστικά ενός περιστατικού εγκυμοσύνης έχουν μεγάλη βαρύτητα στη διάγνωση γενετικών ανωμαλιών και ποια μικρή ή καθόλου;

4.3 Ανάλυση Χαρακτηριστικών

Στην εργασία μου εξέτασα την συνεισφορά και την ευαισθησία των χαρακτηριστικών της κάθε περίπτωσης και πως επηρεάζουν το αποτέλεσμα της πρόβλεψης του συστήματος. Τα χαρακτηριστικά που είχα να εξετάσω ήταν τα εξής:

- **Κεφαλουραίο μήκος (CRL):** Είναι το μήκος του εμβρύου σε εκατοστόμετρα, από την κορυφή της κεφαλής ως το κάτω μέρος των γλουτών. Υπολογίζεται από τον υπέρηχο και μπορεί να προσδιορίσει την ηλικία της κύησης.
- **Αυχενική διαφάνεια (Delta NT):** Η αυχενική διαφάνεια είναι η παρουσία υποδόριου υγρού πίσω από τον αυχένα του εμβρύου κατά τη διάρκεια του πρώτου τριμήνου της εγκυμοσύνης. Η ποσότητα του υγρού που υπάρχει σε αυτό το σημείο διαφέρει από έμβρυο σε έμβρυο. Η αυξημένη ποσότητά του αποτελεί σημαντικό δείκτη για αύξηση της πιθανότητας το έμβρυο να παρουσιάζει κάποια ανωμαλία.

- **Πρωτεΐνη Α (PAPP-A):** Είναι μια ορμόνη που παράγεται στο αίμα της εγκύου, η τιμή της οποίας σχετίζεται με την πιθανότητα για πρόκληση χρωμοσωμικής ανωμαλίας του εμβρύου. Έχει διαπιστωθεί ότι σε γεννήσεις με σύνδρομο Down, η πρωτεΐνη Α είναι χαμηλή.
- **Χοριακή γονιαδοτροπίνη (b-hCG):** Είναι μια ορμόνη που παράγεται από τον πλακούντα κατά τη διάρκεια της εγκυμοσύνης με σκοπό να εμποδίζει την εμμηνόρροια και να προστατεύει την εγκυμοσύνη. Ακόμα, έχει διαπιστωθεί ότι σε γεννήσεις με σύνδρομο Down, η ελεύθερη β-HCG είναι υψηλή.
- **Καρδιακός Σφυγμός Εμβρύου:** Μελέτες έχουν δείξει ότι ο καρδιακός παλμός του εμβρύου σχετίζεται με πιθανές χρωμοσωμικές ανωμαλίες. Διαπιστώθηκε ότι ένας σχετικά χαμηλός καρδιακός παλμός αυξάνει την πιθανότητα εμφάνισης ανωμαλιών.
- **Ρινικό οστό (Nasal Bone):** Είναι τμήμα οστού που βρίσκεται ανάμεσα στις κόγχες των ματιών και εξέχει προς τα έξω. Σύμφωνα με ιατρικές έρευνες η ύπαρξη ρινικού οστού είναι ένα θετικό προμήνυμα για την υγεία του εμβρύου, καθώς το 60% των εμβρύων που δεν έχουν ρινικό οστό έχουν χρωμοσωμική ανωμαλία, ενώ μόνο το 1-3% των χρωμοσωμικά υγιών εμβρύων δεν έχουν ρινικό οστό.
- **Τριγλώχινα ροή (Tricuspid flow):** Είναι ροή αίματος στην τριγλώχινα βαλβίδα. Για την τρισωμία 21 η εκτίμηση της ροής της τριγλώχινας αυξάνει το ποσοστό της ανίχνευσης από 90% σε 95% και μειώνει τα ψευδώς θετικά αποτελέσματα από 3% σε 2,5%.
- **Ροή φλεβώδους πόρου (Ductus Venosus Flow):** Ο φλεβώδης πόρος είναι ένα μικρό αγγείο που συνδέει την ομφαλική φλέβα με την κάτω κοίλη φλέβα και βοηθά στην μεταφορά οξυγονωμένου αίματος στον εγκέφαλο του εμβρύου. Ανωμαλίες στο φλεβώδη πόρο παρατηρούνται στο 80% των εμβρύων με τρισωμία 21, στο 75% των εμβρύων με άλλες χρωμοσωμικές ανωμαλίες και στο 5% των χρωμοσωμικά υγιών εμβρύων.
- **Ηλικία μητέρας:** Είναι ίσως ο πιο κρίσιμος παράγοντας. Σε ηλικία 20 ετών η πιθανότητα παρουσίας χρωμοσωμικής διαταραχής είναι μόλις 1 προς 2000, ενώ στην ηλικία των 40 η πιθανότητα αυξάνεται σε 1 προς 100.
- **Βάρος μητέρας:** Το βάρος της γυναίκας κατά την εγκυμοσύνη μπορεί σε συνδυασμό με άλλους δείκτες και την ηλικία της μητέρας να συντελέσει στον υπολογισμό της πιθανότητας χρωμοσωμικής ανωμαλίας.

- **Σύλληψη:** Ο τρόπος σύλληψης του εμβρύου μπορεί να συντελέσει ως παράγοντας σε μια χρωμοσωμική διαταραχή.
- **Κάπνισμα:** Δυαδικό χαρακτηριστικό που προσδιορίζει κατά πόσον η εγκυμονούσα είναι καπνίστρια ή όχι. Το κάπνισμα μπορεί να συνδέεται με πιθανή γενετική ανωμαλία.
- **Διαβήτης:** Έρευνες έχουν δείξει ότι γυναίκες που είχαν αναπτύξει διαβήτη της κύησης είχαν διπλάσια πιθανότητα να αποκτήσουν παιδιά με χρωμοσωμικές ανωμαλίες. Ο κίνδυνος αφορούσε ιδιαίτερα σε ανωμαλίες των χρωμοσωμάτων του φύλου, ειδικά ανωμαλίες κατά τις οποίες το παιδί είχε ένα επιπλέον χρωμόσωμα X.
- **Κληρονομικότητα (Τρισωμία 21/ Τρισωμία 18/ Τρισωμία 13):** Είναι 3 διαφορετικά δυαδικά χαρακτηριστικά που προσδιορίζουν αν υπήρξε προηγούμενη γέννηση παιδιού από τους συγκεκριμένους γονείς με κάποια συγκεκριμένη τρισωμία.

4.4 Παραδοσιακές Μέθοδοι Αντιμετώπισης

Οι παραδοσιακές μέθοδοι αντιμετώπισης αυτού του προβλήματος στηρίζονταν κυρίως σε επεμβατικές μεθόδους, η πιο γνωστή μέθοδος είναι η αμνιοκέντηση [6], αλλιώς αμνιοπαρακέντηση. Οι επεμβατικές μέθοδοι θεωρούνται πιο ακριβής, ίσως και απόλυτα επιτυχείς, αλλά κρύβουν πολλούς κινδύνους για τη μητέρα και το μωρό. Η εξέταση αυτή συνηθίζεται να γίνεται στις εγκύους ηλικίας άνω των 35 ετών ή αν υπάρχει οικογενειακό ιστορικό χρωμοσωμικών ανωμαλιών. Τα αποτελέσματα της εξέτασης δίνουν πολλές πληροφορίες για την υγεία του εμβρύου. Με την μέθοδο αυτή μπορούν να διαγνωστούν όλες οι χρωμοσωμικές ανωμαλίες με πλήρη αξιοπιστία, αλλά χρειάζονται τουλάχιστον δύο εβδομάδες για να γίνουν γνωστά. Επίσης, υπάρχει ένας μικρός κίνδυνος αποβολής μετά την εξέταση, με πιθανότητα περίπου 1 προς 200, ενώ η πιθανότητα αυτή αυξάνεται αν η εξέταση γίνει πριν από τη 14η εβδομάδα.

Μια άλλη γρηγορότερη μέθοδος, είναι η «FISH» (Fluorescence In Situ Hybridization) [19]. Η εξέταση αυτή μπορεί να εξετάσει μόνο 4 ή 5 διαφορετικά ζεύγη χρωμοσωμάτων από τα 23. Τα πιο συνηθισμένα γενετικά προβλήματα ανιχνεύονται με τη χρήση αυτής της μεθόδου

και τα αποτελέσματα αυτής της εξέτασης είναι διαθέσιμα μέσα σε λίγες ώρες, γι' αυτούς τους λόγους προτιμάται και αυτή η μέθοδος.

Άλλη μια εναλλακτική μέθοδος είναι η λήψη χοριακής λάχνης, CVS (Chorionic Villus Sampling) [20]. Η διαδικασία είναι παρόμοια με την αμνιοπαρακέντηση και κατά την εξέταση αυτή γίνεται λήψη χοριακού ιστού από τον πλακούντα, μέσω του τραχήλου ή του κοιλιακού τοιχώματος και στην συνέχεια αξιολογείται. Τα αποτελέσματα της εξέτασης είναι ακριβή και λαμβάνονται μέσα σε λίγες ημέρες, έχουμε όμως αυξημένο κίνδυνο αποβολής που φτάνει το 2%.

4.5 Σύγκριση και Συζήτηση Μεθόδων

Λόγω της ύπαρξης, έστω και μικρού, κινδύνου αποβολής με την εφαρμογή κάποιας επεμβατικής μεθόδου για την ανάλυση του εμβρυϊκού υλικού, γεννήθηκε η ανάγκη αναζήτησης μη επεμβατικών τρόπων διάγνωσης. Οι γνωστότερες μη επεμβατικές εξετάσεις είναι το υπερηχογράφημα, η μέτρηση της αυχενικής διαφάνειας και η μέτρηση βιοχημικών δεικτών στο αίμα της εγκύου. Από τα αποτελέσματα των εξετάσεων και σε συνδυασμό με την ηλικία της μητέρας, μπορούν να δημιουργηθούν μοντέλα που να «μαθαίνουν» τις διάφορες κατηγορίες έτσι ώστε να εκτιμηθεί ο πιθανός κίνδυνος το έμβρυο να έχει κάποια χρωμοσωμική ανωμαλία.

Υπάρχουν στατιστικές μέθοδοι για την πρόβλεψη αυτών των προβληματικών γενετικών περιπτώσεων, οι οποίοι κατατάσσονται στην κατηγορία των μη επεμβατικών μεθόδων. Μερικές στατιστικές μέθοδοι που χρησιμοποιούνται είναι η ταξινόμηση κατά Bayes, Feature vectors και Γραμμικοί ταξινομητές.

Μια άλλη μη επεμβατική μέθοδος για πρόβλεψη χρωμοσωμικών ανωμαλιών ανιχνεύει τις ανωμαλίες αυτές μέσω εξωκυτταρικών εμβρυϊκών DNA, τα οποία κυκλοφορούν στο αίμα της εγκύου. Σε αυτή την περίπτωση όμως, η ποσότητα εμβρυϊκού DNA μέσα στο αίμα της μητέρας είναι ελάχιστη και η συνύπαρξη μεγάλης ποσότητας μητρικού DNA καθιστά τον διαχωρισμό του εμβρυϊκού DNA από το μητρικό DNA πολύ δύσκολο.

Ο λόγος που αποτρέπει όλες αυτές τις μη επεμβατικές μεθόδους να αντικαταστήσουν τις επεμβατικές οι οποίες είναι επικίνδυνες για τη ζωή της μητέρας και του μωρού, είναι η αναξιοπιστία τους. Τα ποσοστά επιτυχίας στις προβλέψεις τους δεν είναι τόσο υψηλά όσο θα

θέλαμε ούτως ώστε να τις εφαρμόσουμε χωρίς να αφήνουμε κανένα περιθώριο αμφιβολίας όσον αφορά τις προβλέψεις τους, οι οποίες κατ' επέκταση θα ορίσουν το μέλλον της ζωής του εμβρύου.

Κεφάλαιο 5

Πρόταση Εργασίας

5.1 Συνδυασμός Τεχνικών

5.2 Τεχνική 1

5.3 Τεχνική 2

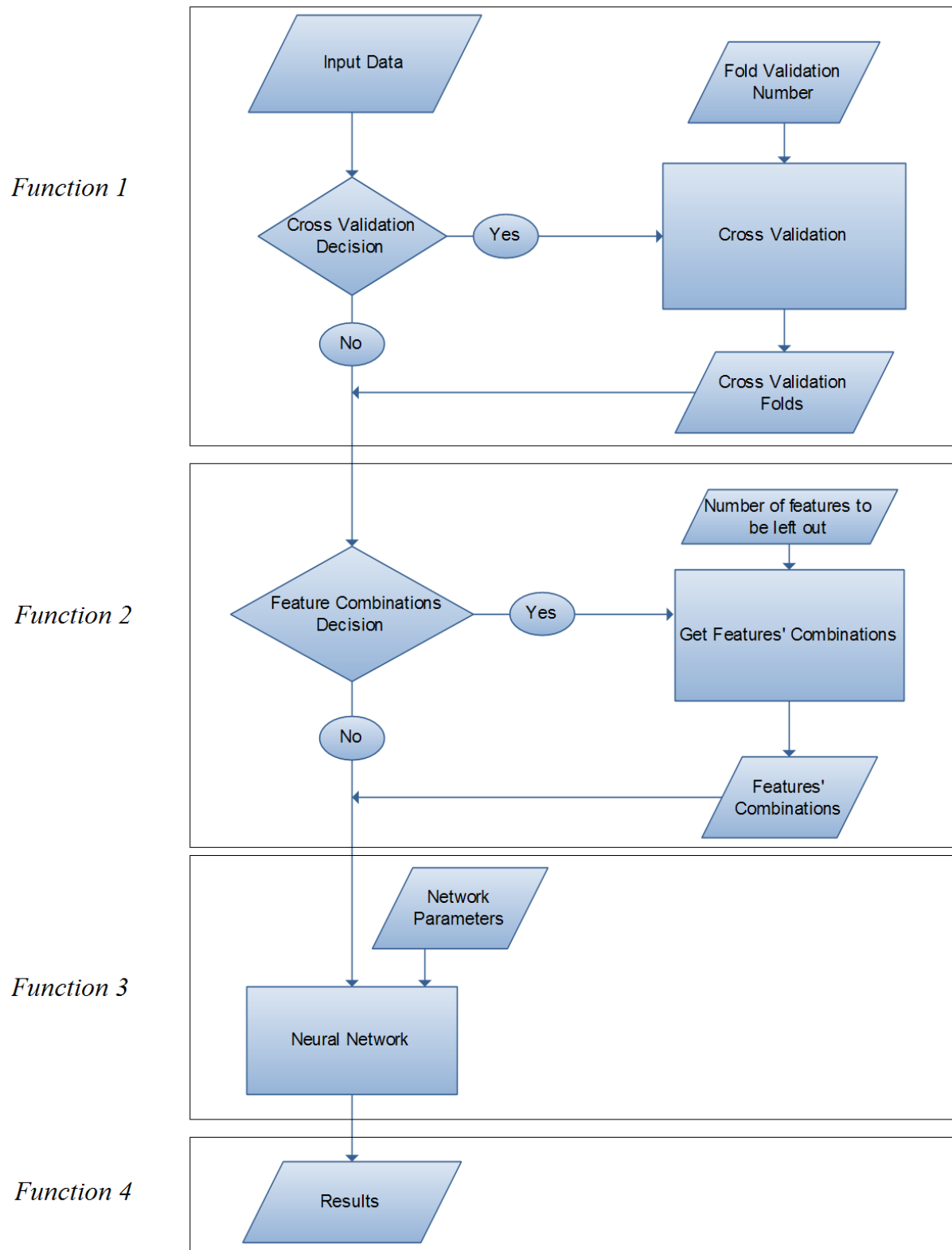
5.4 Τεχνική 3

5.5 Αποτελέσματα

5.1 Συνδυασμός Τεχνικών

Με σκοπό την πρόβλεψη των γενετικών ανωμαλιών χρησιμοποίησα ένα πρόγραμμα που είναι γραμμένο σε Matlab με σκοπό την προσομοίωση νευρωνικών δικτύων. Το πρόγραμμα αυτό δημιουργήθηκε από τον ερευνητή Αντρέα Νεοκλέους. Το πρόγραμμα υποστήριζε διάφορες συναρτήσεις που υλοποιούν η καθεμία μια διαφορετική διεργασία ούτως ώστε να μοντελοποιούν πολλά διαφορετικά νευρωνικά δίκτυα, με διαφορετικές αρχιτεκτονικές και συνδυασμούς παραμέτρων. Όλες οι τεχνικές μπορούν να συνεργάζονται μαζί δημιουργώντας πολλαπλές επιλογές και δυνατότητες για δημιουργία πολλών νευρωνικών δικτύων, δοκιμή διαφόρων παραμέτρων αλλά και επεξεργασία των δεδομένων εισόδου. Έχουμε την επιλογή κάθε μια από αυτές να την χρησιμοποιήσουμε ή όχι. Έτσι ο αλγόριθμος τρέχοντας τον με τις κατάλληλες εισόδους, δημιουργεί από μόνος του συνδυασμούς δεδομένων σύμφωνα με το Cross Validation, στη συνέχεια μπορεί να επιλέξει να αφήσει κάποια χαρακτηριστικά εκτός εκπαίδευσης του δικτύου, ακολούθως να κατασκευάσει με πολλές διαφορετικές παραμέτρους διάφορα μοντέλα, να υπολογίσει τα αποτελέσματα, τα ποσοστά επιτυχίας, και να τα

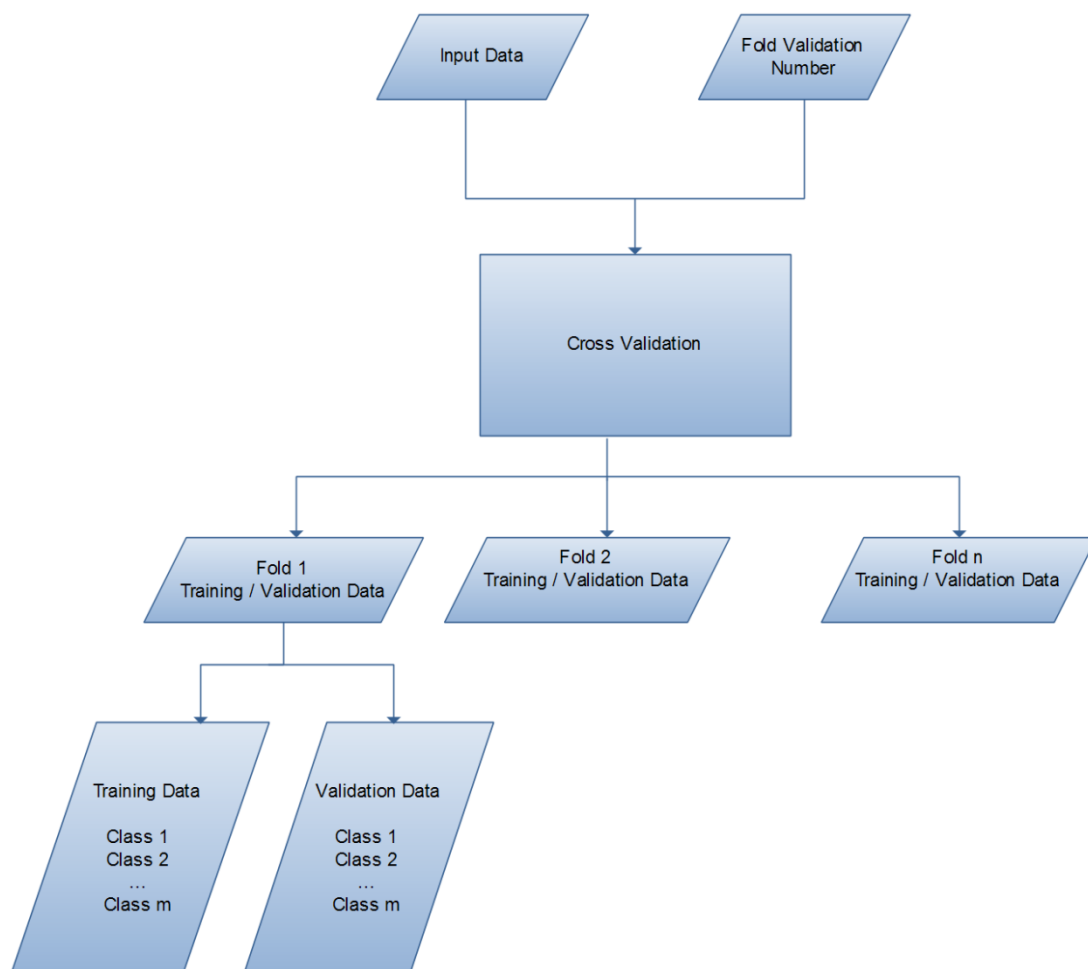
παρουσιάσει. Ο αλγόριθμος επίσης έχει την επιλογή να αποθηκεύει τα αποτελέσματα αυτά σε ένα αρχείο τύπου Excel. Η ολοκληρωμένη διαδικασία όλων των λειτουργιών του αλγορίθμου φαίνεται στο πιο κάτω διάγραμμα ροής δεδομένων (σχήμα 5.1.1).



Σχήμα 5.1.1: Απεικόνιση διαδικασίας συνδυασμού τεχνικών.

5.2 Τεχνική 1

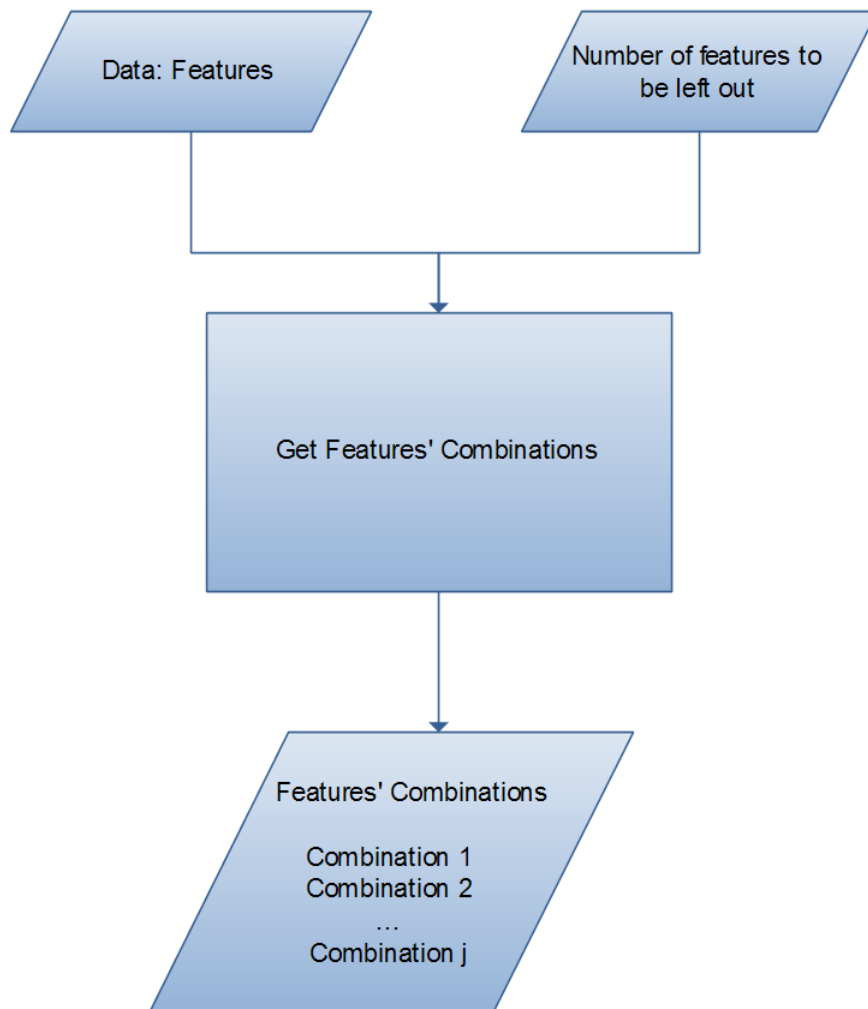
Η πρώτη τεχνική που υποστηρίζει ο κώδικας είναι η επιλογή του ποσοστού του Cross-Validation και η αυτόματη εναλλαγή των δεδομένων για δεδομένα εκπαίδευσης και επαλήθευσης αντίστοιχα. Συγκεκριμένα η συνάρτηση αυτή δέχεται σαν είσοδο τα δεδομένα, μαζί και με τον αριθμό του Cross Validation, ένας αριθμός που συμβολίζει τον αριθμών των διαφορετικών συνόλων για δεδομένα εκπαίδευσης και επαλήθευσης. Ακολούθως η συνάρτηση αυτή επιστρέφει όλα τα διαφορετικά σύνολα για τα δύο σετ δεδομένων, εκπαίδευσης και επαλήθευσης. Η διαδικασία αυτή φαίνεται στο πιο κάτω διάγραμμα ροής δεδομένων (σχήμα 5.2.1).



Σχήμα 5.2.1: Απεικόνιση διαδικασίας Cross Validation.

5.3 Τεχνική 2

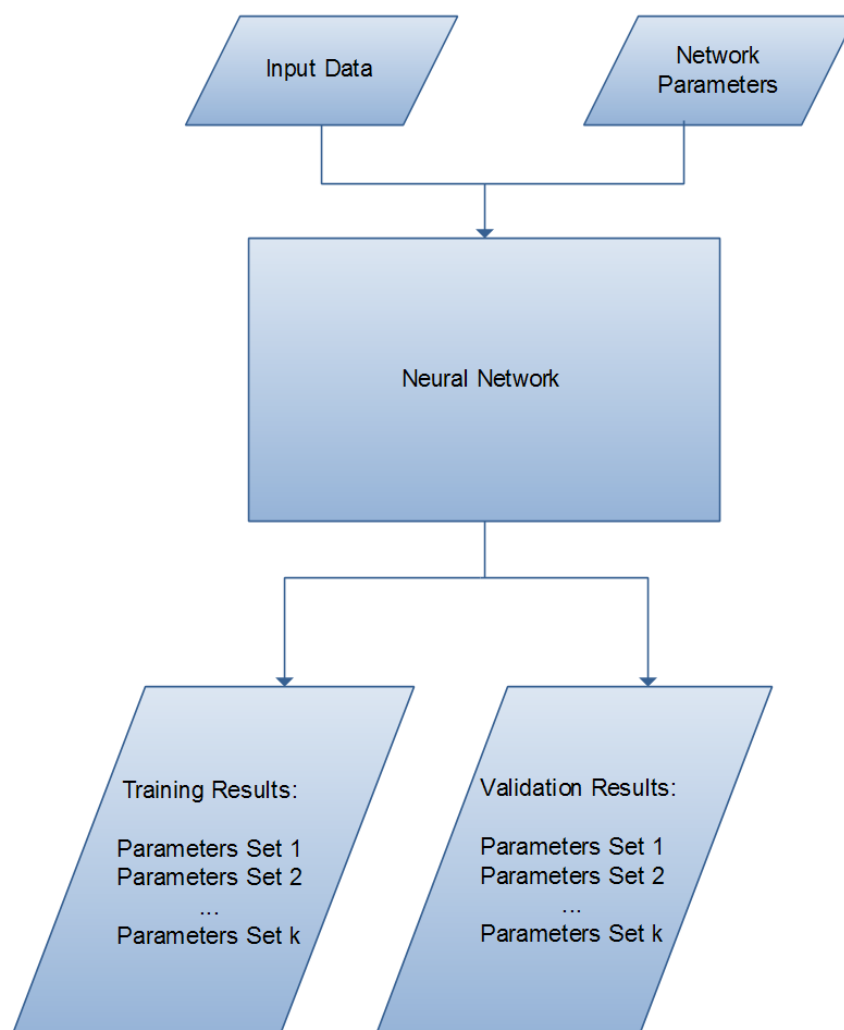
Η δεύτερη συνάρτηση του προγράμματος έχει την ικανότητα να παράγει συνδυασμούς χαρακτηριστικών εισόδου στο νευρωνικό δίκτυο. Πιο συγκεκριμένα η συνάρτηση λαμβάνει σαν είσοδο ένα αριθμό ο οποίος θα συμβολίζει τον αριθμό των χαρακτηριστικών που θα αφήνει έξω σε κάθε κάλεσμα της, καθώς και το σύνολο των χαρακτηριστικών. Στη συνέχεια θα υπολογίζει και θα επιστρέφει όλους τους δυνατούς συνδυασμούς χαρακτηριστικών εισόδου με το υπόλοιπο του αριθμού που θα επιλέξουμε να αφήσουμε στα δεδομένα μας. Ο αριθμός των χαρακτηριστικών που θα αφήνει έξω με το κάλεσμα της είναι μεταβλητός και τον επιλέγω εγώ. Η διαδικασία αυτή φαίνεται επίσης στο πιο κάτω διάγραμμα ροής δεδομένων (σχήμα 5.3.1).



Σχήμα 5.3.1: Απεικόνιση διαδικασίας Get Features.

5.4 Τεχνική 3

Η επόμενη τεχνική που υποστηρίζει ο κώδικας, και πιο σημαντική είναι η υλοποίηση των μοντέλων νευρωνικών δικτύων. Σε αυτή τη συνάρτηση δίνονται σαν είσοδοι τα σετ δεδομένων εκπαίδευσης και επαλήθευσης καθώς και το εύρος τιμών που θα κινηθούν οι παράμετροι που θα χρησιμοποιηθούν στα διάφορα μοντέλα που θα υλοποιήσω. Ακολούθως η συνάρτηση κατασκευάζει όλους τους συνδυασμούς νευρωνικών δικτύων, βασισμένη στις διάφορες παραμέτρους που δόθηκαν σαν είσοδος, εκπαιδεύοντας και επαληθεύοντας τα με τα δεδομένα εκπαίδευσης και επαλήθευσης αντίστοιχα. Στη συνέχεια βασισμένο στα επιθυμητά αποτελέσματα και στα αποτελέσματα εξόδου των νευρωνικών δικτύων, υπολογίζει τα ποσοστά επιτυχίας για όλους τους συνδυασμούς που κατασκεύασε, αποθηκεύει και επιστρέφει τα αποτελέσματα σε πίνακες. Η διαδικασία της κατασκευής των νευρωνικών δικτύων φαίνεται εικονικά στο πιο κάτω διάγραμμα ροής δεδομένων (σχήμα 5.4.1).



Σχήμα 5.4.1: Απεικόνιση διαδικασίας Create Neural Network.

5.5 Αποτελέσματα

Χρησιμοποιήθηκαν 2 βάσεις δεδομένων, η μία με raw data και η άλλη με MOM. Οι βάσεις δεδομένων είχαν δεδομένα με 2 κατηγορίες, αυτές των φυσιολογικών εγκυμοσύνων και αυτών με τρισωμία 21. Δημιουργήθηκαν μοντέλα με εναλλαγές σε παραμέτρους όπως αριθμός κρυφών επιπέδων νευρώνων και νευρώνων, εποχών αλλά και συνδυασμών χαρακτηριστικών εισόδου. Πιο κάτω παρουσιάζονται τα αποτελέσματα από τα διάφορα μοντέλα που υλοποιήθηκαν.

Πείραμα 1

Στο πρώτο πείραμα δημιουργούνται διάφορα νευρωνικά δίκτυα με ένα κρυφό επίπεδο στα οποία γίνεται ένας συνδυασμός των παραμέτρων εναλλάσσοντας τον αριθμό των εποχών και των νευρώνων του κρυφού επιπέδου. Οι εποχές κινούνται από τις 200 μέχρι τις 1200 αυξανόμενες κατά 200 κάθε φορά, ενώ ο αριθμός των νευρώνων του κρυφού επιπέδου κυμαίνεται από τους 5 μέχρι τους 20, αυξανόμενοι κατά 5 κάθε φορά. Η συνάρτηση ενεργοποίησης του επιπέδου εισόδου είναι η $\tanh$ και η συνάρτηση ενεργοποίησης του κρυφού επιπέδου είναι η purelin . Ο ρυθμός μάθησης και η ορμή είναι επίσης σταθερά στο 0.1. Ο αλγόριθμος χώρισε τυχαία τα δεδομένα σε δεδομένα εκπαίδευσης και δεδομένα επαλήθευσης με ποσοστά 67% και 33% του συνόλου, αντίστοιχα, βασισμένος στην επιλογή για σύνολο σετ που έδωσε ο χρήστης.

Training Results					Validation Results				
Model ID	Epochs	Neurons	T21	Normal	Model ID	Epochs	Neurons	T21	Normal
newff_1	200	5	92.68 %	94.32 %	newff_1	200	5	80.95 %	93.64 %
newff_2	400	5	91.46 %	97.67 %	newff_2	400	5	80.95 %	97.49 %
newff_3	600	5	91.46 %	98.48 %	newff_3	600	5	78.57 %	98.30 %
newff_4	800	5	91.46 %	98.55 %	newff_4	800	5	78.57 %	98.33 %
newff_5	1000	5	91.46 %	98.62 %	newff_5	1000	5	78.57 %	98.36 %
newff_6	1200	5	91.46 %	98.68 %	newff_6	1200	5	78.57 %	98.40 %
newff_7	200	10	92.68 %	94.18 %	newff_7	200	10	73.81 %	94.37 %
newff_8	400	10	95.12 %	94.69 %	newff_8	400	10	76.19 %	94.60 %
newff_9	600	10	85.37 %	99.58 %	newff_9	600	10	78.57 %	95.99 %
newff_10	800	10	86.59 %	99.58 %	newff_10	800	10	78.57 %	96.33 %
newff_11	1000	10	86.59 %	99.58 %	newff_11	1000	10	78.57 %	96.37 %
newff_12	1200	10	86.59 %	99.58 %	newff_12	1200	10	78.57 %	96.48 %
newff_13	200	15	96.34 %	94.79 %	newff_13	200	15	88.10 %	94.77 %
newff_14	400	15	97.56 %	94.99 %	newff_14	400	15	85.71 %	94.46 %
newff_15	600	15	98.78 %	94.97 %	newff_15	600	15	85.71 %	94.91 %
newff_16	800	15	98.78 %	94.96 %	newff_16	800	15	85.71 %	94.97 %
newff_17	1000	15	98.78 %	94.96 %	newff_17	1000	15	88.10 %	94.99 %
newff_18	1200	15	98.78 %	94.93 %	newff_18	1200	15	88.10 %	95.00 %
newff_19	200	20	100.00 %	94.92 %	newff_19	200	20	71.43 %	94.85 %
newff_20	400	20	86.59 %	99.59 %	newff_20	400	20	64.29 %	94.90 %
newff_21	600	20	89.02 %	99.59 %	newff_21	600	20	66.67 %	94.83 %
newff_22	800	20	89.02 %	99.59 %	newff_22	800	20	64.29 %	94.29 %
newff_23	1000	20	90.24 %	99.59 %	newff_23	1000	20	54.76 %	94.94 %
newff_24	1200	20	93.90 %	99.59 %	newff_24	1200	20	61.90 %	94.42 %

Πείραμα 2

Στο δεύτερο πείραμα δημιουργούνται πολλά νευρωνικά δίκτυα με δύο κρυφά επίπεδα στα οποία πάλι αλλάζει ο αριθμός εποχών εκπαίδευσης και ο αριθμός των νευρώνων των 2 επιπέδων. Πιο συγκεκριμένα ο αριθμός των εποχών κινείται από τις 100 μέχρι τις 800 ανά 100, ενώ οι νευρώνες στα 2 κρυφά επίπεδα κινούνται από 5 σε 10. Ο αλγόριθμος κάνει όλους τους συνδυασμούς με βάση αυτές τις παραμέτρους, ενώ η συνάρτηση ενεργοποίησης του επιπέδου εισόδου είναι η tansig και η συνάρτηση ενεργοποίησης των 2 κρυφών επιπέδων είναι η purelin, για όλα τα μοντέλα.

Training Results:					Validation Results:				
Model ID	Epochs	Neurons	T21	Normal	Model ID	Epochs	Neurons	T21	Normal
newff_1	100	[5,5]	96,34%	92,81%	newff_1	100	[5,5]	73,81%	94,99%
newff_2	200	[5,5]	96,34%	94,78%	newff_2	200	[5,5]	78,57%	94,64%
newff_3	300	[5,5]	96,34%	93,99%	newff_3	300	[5,5]	78,57%	94,01%
newff_4	400	[5,5]	96,34%	94,05%	newff_4	400	[5,5]	76,19%	94,00%
newff_5	500	[5,5]	96,34%	94,44%	newff_5	500	[5,5]	76,19%	94,35%
newff_6	600	[5,5]	96,34%	94,45%	newff_6	600	[5,5]	76,19%	94,35%
newff_7	700	[5,5]	95,12%	94,52%	newff_7	700	[5,5]	76,19%	94,53%
newff_8	800	[5,5]	95,12%	94,52%	newff_8	800	[5,5]	76,19%	94,54%
newff_9	100	[10,10]	98,78%	94,75%	newff_9	100	[10,10]	88,10%	94,97%
newff_10	200	[10,10]	98,78%	94,95%	newff_10	200	[10,10]	90,48%	94,98%
newff_11	300	[10,10]	98,78%	95,00%	newff_11	300	[10,10]	90,48%	94,70%
newff_12	400	[10,10]	98,78%	94,72%	newff_12	400	[10,10]	90,48%	94,76%
newff_13	500	[10,10]	98,78%	94,76%	newff_13	500	[10,10]	90,48%	94,78%
newff_14	600	[10,10]	98,78%	94,77%	newff_14	600	[10,10]	90,48%	94,75%
newff_15	700	[10,10]	98,78%	94,79%	newff_15	700	[10,10]	90,48%	94,76%
newff_16	800	[10,10]	98,78%	94,80%	newff_16	800	[10,10]	90,48%	94,76%

Πείραμα 3

Το επόμενο πείραμα έχει να κάνει με νευρωνικά δίκτυα 2 κρυφών επιπέδων. Ο αριθμός των εποχών παραμένει σταθερός στις 200 σε όλα τα μοντέλα που υλοποιεί ο αλγόριθμος, ενώ αλλάζει σε κάθε μοντέλο ο αριθμός των νευρώνων των δύο κρυφών επιπέδων. Πιο συγκεκριμένα ο αριθμός αυτός ξεκινά από το 5 και φτάνει μέχρι το 50 αυξάνοντας κατά 5 τους νευρώνες κάθε φορά. Η συνάρτηση ενεργοποίησης του επιπέδου εισόδου είναι η tansig και η συνάρτηση ενεργοποίησης των 2 κρυφών επιπέδων είναι η purelin και παραμένουν οι ίδιες σε όλα τα μοντέλα. Στο πείραμα αυτό τα σύνολα εναλλάσσονται 3 φορές ούτως ώστε όλα τα δεδομένα που δημιουργεί η διεργασία Cross Validation να βρεθούν σε κάποια φάση σαν δεδομένα επαλήθευσης.

Fold: 1							
Training Results:				Validation Results:			
Model ID	Neurons	T21	Normal	Model ID	Neurons	T21	Normal
newff_1	[5,5]	96,34%	94,78%	newff_1	[5,5]	78,57%	94,64%
newff_2	[10,10]	98,78%	94,95%	newff_2	[10,10]	90,48%	94,98%
newff_3	[15,15]	95,12%	94,77%	newff_3	[15,15]	80,95%	94,27%
newff_4	[20,20]	93,90%	94,78%	newff_4	[20,20]	85,71%	94,85%
newff_5	[25,25]	100,00%	94,98%	newff_5	[25,25]	73,81%	94,95%
newff_6	[30,30]	98,78%	94,72%	newff_6	[30,30]	78,57%	94,70%
newff_7	[35,35]	98,78%	94,61%	newff_7	[35,35]	45,24%	94,88%
newff_8	[40,40]	98,78%	94,72%	newff_8	[40,40]	78,57%	94,91%
newff_9	[45,45]	100,00%	94,91%	newff_9	[45,45]	78,57%	94,97%
newff_10	[50,50]	98,78%	94,96%	newff_10	[50,50]	76,19%	94,86%
Fold: 2							
Training Results:				Validation Results:			
Model ID	Neurons	T21	Normal	Model ID	Neurons	T21	Normal
newff_1	[5,5]	100,00%	85,78%	newff_1	[5,5]	88,10%	85,44%
newff_2	[10,10]	100,00%	94,93%	newff_2	[10,10]	83,33%	94,97%
newff_3	[15,15]	96,34%	94,89%	newff_3	[15,15]	85,71%	94,81%
newff_4	[20,20]	98,78%	94,13%	newff_4	[20,20]	71,43%	93,62%
newff_5	[25,25]	97,56%	94,00%	newff_5	[25,25]	80,95%	93,87%
newff_6	[30,30]	97,56%	94,84%	newff_6	[30,30]	88,10%	94,89%
newff_7	[35,35]	100,00%	94,91%	newff_7	[35,35]	76,19%	94,87%
newff_8	[40,40]	100,00%	94,74%	newff_8	[40,40]	92,86%	94,71%
newff_9	[45,45]	100,00%	94,85%	newff_9	[45,45]	69,05%	94,86%
newff_10	[50,50]	100,00%	94,79%	newff_10	[50,50]	76,19%	94,86%
Fold: 3							
Training Results:				Validation Results:			
Model ID	Neurons	T21	Normal	Model ID	Neurons	T21	Normal
newff_1	[5,5]	96,34%	83,74%	newff_1	[5,5]	80,95%	84,06%
newff_2	[10,10]	100,00%	94,82%	newff_2	[10,10]	97,62%	94,72%
newff_3	[15,15]	96,34%	87,49%	newff_3	[15,15]	88,10%	87,71%
newff_4	[20,20]	97,56%	94,72%	newff_4	[20,20]	73,81%	94,97%
newff_5	[25,25]	100,00%	94,99%	newff_5	[25,25]	83,33%	94,78%
newff_6	[30,30]	98,78%	94,88%	newff_6	[30,30]	80,95%	94,82%
newff_7	[35,35]	97,56%	94,97%	newff_7	[35,35]	57,14%	94,78%
newff_8	[40,40]	97,56%	94,89%	newff_8	[40,40]	78,57%	94,79%
newff_9	[45,45]	100,00%	94,80%	newff_9	[45,45]	61,90%	94,90%
newff_10	[50,50]	98,78%	94,61%	newff_10	[50,50]	80,95%	94,80%

Πείραμα 4

Σε αυτό το πείραμα δημιουργούνται νευρωνικά δίκτυα ενός κρυφού επιπέδου. Ο αριθμός των εποχών παραμένει σταθερός στις 200 σε όλα τα μοντέλα που υλοποιεί ο αλγόριθμος, και εναλλάσσονται οι νευρώνες του κρυφού επιπέδου ξεκινώντας από 15 και φτάνοντας μέχρι τους 70 αυξανόμενοι κατά 5 σε κάθε διαφορετικό μοντέλο. Η συνάρτηση ενεργοποίησης του επιπέδου εισόδου είναι η tansig και η συνάρτηση ενεργοποίησης του κρυφού επιπέδου είναι η purelin. Και σε αυτό το πείραμα τα σύνολα εναλλάσσονται 3 φορές ούτως ώστε όλα τα δεδομένα που δημιουργεί η διεργασία Cross Validation να βρεθούν σε κάποια φάση σαν δεδομένα επαλήθευσης.

Fold: 1							
Training Results:				Validation Results:			
Model ID	Neurons	T21	Normal	Model ID	Neurons	T21	Normal
newff_1	15	97,56%	94,20%	newff_3	15	73,81%	93,69%
newff_2	20	98,78%	94,59%	newff_4	20	54,76%	94,83%
newff_3	25	98,78%	94,80%	newff_5	25	78,57%	94,91%
newff_4	30	98,78%	94,86%	newff_6	30	78,57%	94,87%
newff_5	35	98,78%	94,97%	newff_7	35	66,67%	94,86%
newff_6	40	98,78%	94,87%	newff_8	40	76,19%	94,16%
newff_7	45	100,00%	94,70%	newff_9	45	80,95%	94,97%
newff_8	50	98,78%	94,84%	newff_10	50	59,52%	94,84%
newff_9	55	100,00%	94,79%	newff_11	55	71,43%	94,86%
newff_10	60	100,00%	94,89%	newff_12	60	54,76%	94,86%
newff_11	65	100,00%	94,84%	newff_13	65	59,52%	94,99%
newff_12	70	100,00%	94,41%	newff_14	70	66,67%	94,80%
Fold: 2							
Training Results:				Validation Results:			
Model ID	Neurons	T21	Normal	Model ID	Neurons	T21	Normal
newff_1	15	97,56%	94,89%	newff_3	15	71,43%	94,92%
newff_2	20	97,56%	94,29%	newff_4	20	83,33%	94,88%
newff_3	25	95,12%	93,92%	newff_5	25	80,95%	94,81%
newff_4	30	100,00%	94,81%	newff_6	30	69,05%	94,80%
newff_5	35	100,00%	94,76%	newff_7	35	83,33%	94,69%
newff_6	40	100,00%	94,90%	newff_8	40	80,95%	94,98%
newff_7	45	100,00%	94,97%	newff_9	45	73,81%	94,97%
newff_8	50	100,00%	94,90%	newff_10	50	59,52%	94,71%
newff_9	55	100,00%	94,83%	newff_11	55	50,00%	94,74%
newff_10	60	100,00%	94,96%	newff_12	60	64,29%	94,98%
newff_11	65	100,00%	94,42%	newff_13	65	50,00%	94,99%
newff_12	70	100,00%	94,89%	newff_14	70	61,90%	94,89%

Fold: 3							
Training Results:				Validation Results:			
Model ID	Neurons	T21	Normal	Model ID	Neurons	T21	Normal
newff_1	15	98,78%	94,77%	newff_3	15	85,71%	94,67%
newff_2	20	95,12%	94,37%	newff_4	20	76,19%	94,08%
newff_3	25	100,00%	94,84%	newff_5	25	85,71%	94,80%
newff_4	30	98,78%	94,93%	newff_6	30	83,33%	94,72%
newff_5	35	100,00%	94,82%	newff_7	35	78,57%	94,98%
newff_6	40	98,78%	94,98%	newff_8	40	78,57%	94,71%
newff_7	45	96,34%	95,00%	newff_9	45	85,71%	94,86%
newff_8	50	100,00%	94,95%	newff_10	50	78,57%	94,90%
newff_9	55	100,00%	94,92%	newff_11	55	78,57%	94,96%
newff_10	60	100,00%	94,81%	newff_12	60	83,33%	94,91%
newff_11	65	100,00%	94,79%	newff_13	65	40,48%	94,81%
newff_12	70	100,00%	94,86%	newff_14	70	76,19%	94,98%

Πείραμα 5

Σε αυτό το πείραμα χρησιμοποιήσαμε τη δυνατότητα που προσφέρει ο αλγόριθμος για δημιουργία συνδυασμών των χαρακτηριστικών εισόδου, σύμφωνα με επιλογή που διαλέγει ο χρήστης όπως το ποια χαρακτηριστικά θα χρησιμοποιηθούν και τον αριθμό που θα αφήνονται εκτός σε κάθε κατασκευή νευρωνικού δικτύου, εκπαίδευσης και επαλήθευσης του. Σε αυτό το πείραμα επιλέξαμε να χρησιμοποιήσουμε τα πρώτα 9 χαρακτηριστικά από τα 18 που έχει η βάση δεδομένων μας. Τα χαρακτηριστικά σε συντομογραφία είναι τα εξής: CRL, NT, PAPP-A, b-hCG, Heart rate, Nasal Bone, Tricuspid flow, Ductus venosus, PI, MA. Στην αρχή τρέξαμε ένα φάσμα μοντέλων αλλάζοντας μόνο τους νευρώνες του κρυφού επιπέδου, οι οποίοι μετακινήθηκαν από 5 μέχρι 25, ανά 5. Ο αριθμός εποχών εκπαίδευσης παρέμενε σταθερός στις 200 όπως και οι συναρτήσεις ενεργοποίησης που ήταν tansig και purelin για επίπεδο εισόδου και κρυφού επιπέδου αντίστοιχα. Από τα αποτελέσματα επιλέχθηκε το μοντέλο με τα σχετικά καλύτερα αποτελέσματα, το οποίο ήταν αυτό με τους 10 νευρώνες στο κρυφό επίπεδο για να χρησιμοποιηθεί στη συνέχεια. Ακολούθως εφαρμόσαμε έλεγχο στα χαρακτηριστικά κάνοντας στο σύνολο 8 διαφορετικά σύνολα μοντέλων, στα οποία κάθε φορά αφήναμε εκτός της διαδικασίας ένα αριθμό χαρακτηριστικών. Για κάθε διαφορετικό σύνολο το πρόγραμμα έκανε όλους τους πιθανούς συνδυασμούς χαρακτηριστικών αφήνοντας το συγκεκριμένο αριθμό χαρακτηριστικών έξω, δημιουργώντας στο σύνολο 510 μοντέλα και παρουσιάζοντας τα αποτελέσματα, δείχνοντας μάλιστα για το κάθε ένα ποια χαρακτηριστικά αγνοήθηκαν από το μοντέλο.

Training Results				Validation Results			
Model ID	Neurons	T21	Normal	Model ID	Neurons	T21	Normal
'newff_1'	5	96,34%	93,97%	'newff_1'	5	92,86%	93,78%
'newff_2'	10	98,78%	94,72%	'newff_2'	10	83,33%	94,67%
'newff_3'	15	97,56%	93,93%	'newff_3'	15	76,19%	94,10%
'newff_4'	20	98,78%	94,94%	'newff_4'	20	76,19%	94,95%
'newff_5'	25	98,78%	94,91%	'newff_5'	25	83,33%	94,09%

Test 1 (1 feature left out):							
Training Results:				Validation Results:			
Model ID	Features left out	T21	Normal	Model ID	Features left out	T21	Normal
'newff_1'	1	96,34%	94,97%	'newff_1'	1	92,86%	94,98%
'newff_2'	2	89,02%	94,66%	'newff_2'	2	83,33%	94,30%
'newff_3'	3	92,68%	94,48%	'newff_3'	3	95,24%	94,59%
'newff_4'	4	96,34%	94,91%	'newff_4'	4	92,86%	94,61%
'newff_5'	5	98,78%	94,68%	'newff_5'	5	92,86%	94,90%
'newff_6'	6	95,12%	80,96%	'newff_6'	6	90,48%	80,79%
'newff_7'	7	96,34%	94,51%	'newff_7'	7	97,62%	94,73%
'newff_8'	8	95,12%	94,96%	'newff_8'	8	92,86%	94,83%
'newff_9'	9	90,24%	94,87%	'newff_9'	9	83,33%	94,94%
Test 2 (2 features left out):							
Training Results:				Validation Results:			
Model ID	Features left out	T21	Normal	Model ID	Features left out	T21	Normal
'newff_1'	[1,2]	89,02%	94,47%	'newff_1'	[1,2]	85,71%	94,95%
'newff_2'	[1,3]	84,15%	94,91%	'newff_2'	[1,3]	78,57%	94,95%
'newff_3'	[1,4]	93,90%	94,87%	'newff_3'	[1,4]	83,33%	93,71%
'newff_4'	[1,5]	96,34%	94,40%	'newff_4'	[1,5]	95,24%	94,67%
'newff_5'	[1,6]	91,46%	98,80%	'newff_5'	[1,6]	78,57%	98,62%
'newff_6'	[1,7]	95,12%	94,98%	'newff_6'	[1,7]	90,48%	94,97%
'newff_7'	[1,8]	97,56%	93,31%	'newff_7'	[1,8]	95,24%	93,67%
'newff_8'	[1,9]	97,56%	92,75%	'newff_8'	[1,9]	95,24%	93,67%
'newff_9'	[2,3]	82,93%	94,35%	'newff_9'	[2,3]	83,33%	94,27%
'newff_10'	[2,4]	85,37%	94,91%	'newff_10'	[2,4]	80,95%	94,94%
'newff_11'	[2,5]	92,68%	94,97%	'newff_11'	[2,5]	90,48%	94,96%
'newff_12'	[2,6]	93,90%	94,93%	'newff_12'	[2,6]	83,33%	94,88%
'newff_13'	[2,7]	86,59%	94,81%	'newff_13'	[2,7]	92,86%	93,31%
'newff_14'	[2,8]	86,59%	94,83%	'newff_14'	[2,8]	76,19%	94,73%
'newff_15'	[2,9]	89,02%	94,50%	'newff_15'	[2,9]	83,33%	94,91%
'newff_16'	[3,4]	96,34%	94,03%	'newff_16'	[3,4]	92,86%	94,35%
'newff_17'	[3,5]	93,90%	94,79%	'newff_17'	[3,5]	95,24%	94,36%

'newff_18'	[3,6]	90,24%	94,84%		'newff_18'	[3,6]	95,24%	94,84%
'newff_19'	[3,7]	96,34%	94,82%		'newff_19'	[3,7]	92,86%	94,81%
'newff_20'	[3,8]	93,90%	94,52%		'newff_20'	[3,8]	90,48%	94,89%
'newff_21'	[3,9]	82,93%	98,55%		'newff_21'	[3,9]	76,19%	98,51%
'newff_22'	[4,5]	93,90%	60,36%		'newff_22'	[4,5]	83,33%	60,42%
'newff_23'	[4,6]	90,24%	94,88%		'newff_23'	[4,6]	90,48%	94,71%
'newff_24'	[4,7]	93,90%	94,91%		'newff_24'	[4,7]	95,24%	94,77%
'newff_25'	[4,8]	96,34%	94,76%		'newff_25'	[4,8]	88,10%	94,83%
'newff_26'	[4,9]	81,71%	99,22%		'newff_26'	[4,9]	69,05%	99,22%
'newff_27'	[5,6]	93,90%	94,97%		'newff_27'	[5,6]	83,33%	94,79%
'newff_28'	[5,7]	96,34%	91,55%		'newff_28'	[5,7]	90,48%	91,80%
'newff_29'	[5,8]	95,12%	94,23%		'newff_29'	[5,8]	90,48%	94,93%
'newff_30'	[5,9]	96,34%	94,37%		'newff_30'	[5,9]	95,24%	94,75%
'newff_31'	[6,7]	93,90%	94,91%		'newff_31'	[6,7]	88,10%	94,85%
'newff_32'	[6,8]	93,90%	94,97%		'newff_32'	[6,8]	95,24%	94,90%
'newff_33'	[6,9]	95,12%	94,88%		'newff_33'	[6,9]	85,71%	94,77%
'newff_34'	[7,8]	96,34%	94,86%		'newff_34'	[7,8]	88,10%	94,83%
'newff_35'	[7,9]	91,46%	35,30%		'newff_35'	[7,9]	92,86%	35,23%
'newff_36'	[8,9]	90,24%	91,84%		'newff_36'	[8,9]	76,19%	92,37%
Test 3 (3 features left out):								
Training Results:					Validation Results:			
Model ID	Features left out	T21	Normal		Model ID	Features left out	T21	Normal
'newff_1'	[1,2,3]	84,15%	94,73%		'newff_1'	[1,2,3]	80,95%	94,77%
'newff_2'	[1,2,4]	82,93%	94,70%		'newff_2'	[1,2,4]	76,19%	94,89%
'newff_3'	[1,2,5]	92,68%	92,23%		'newff_3'	[1,2,5]	90,48%	92,30%
'newff_4'	[1,2,6]	89,02%	94,96%		'newff_4'	[1,2,6]	83,33%	94,83%
'newff_5'	[1,2,7]	84,15%	93,94%		'newff_5'	[1,2,7]	80,95%	93,78%
'newff_6'	[1,2,8]	85,37%	94,98%		'newff_6'	[1,2,8]	76,19%	94,27%
'newff_7'	[1,2,9]	71,95%	98,92%		'newff_7'	[1,2,9]	57,14%	98,76%
'newff_8'	[1,3,4]	93,90%	94,82%		'newff_8'	[1,3,4]	88,10%	94,03%
'newff_9'	[1,3,5]	98,78%	94,38%		'newff_9'	[1,3,5]	97,62%	94,45%
'newff_10'	[1,3,6]	93,90%	94,91%		'newff_10'	[1,3,6]	88,10%	94,93%
'newff_11'	[1,3,7]	98,78%	94,89%		'newff_11'	[1,3,7]	90,48%	94,96%
'newff_12'	[1,3,8]	84,15%	94,83%		'newff_12'	[1,3,8]	80,95%	94,85%
'newff_13'	[1,3,9]	95,12%	92,50%		'newff_13'	[1,3,9]	90,48%	92,81%
'newff_14'	[1,4,5]	93,90%	94,54%		'newff_14'	[1,4,5]	88,10%	94,92%
'newff_15'	[1,4,6]	87,80%	94,17%		'newff_15'	[1,4,6]	78,57%	94,45%
'newff_16'	[1,4,7]	95,12%	94,94%		'newff_16'	[1,4,7]	90,48%	94,92%
'newff_17'	[1,4,8]	96,34%	93,98%		'newff_17'	[1,4,8]	66,67%	94,29%
'newff_18'	[1,4,9]	85,37%	94,57%		'newff_18'	[1,4,9]	88,10%	94,80%
'newff_19'	[1,5,6]	92,68%	94,53%		'newff_19'	[1,5,6]	83,33%	94,67%
'newff_20'	[1,5,7]	92,68%	94,86%		'newff_20'	[1,5,7]	88,10%	94,89%
'newff_21'	[1,5,8]	93,90%	94,70%		'newff_21'	[1,5,8]	88,10%	94,89%
'newff_22'	[1,5,9]	97,56%	94,74%		'newff_22'	[1,5,9]	95,24%	94,27%
'newff_23'	[1,6,7]	93,90%	94,89%		'newff_23'	[1,6,7]	90,48%	94,54%
'newff_24'	[1,6,8]	96,34%	74,37%		'newff_24'	[1,6,8]	97,62%	74,06%

'newff_25'	[1,6,9]	89,02%	94,61%	'newff_25'	[1,6,9]	85,71%	94,56%
'newff_26'	[1,7,8]	96,34%	94,71%	'newff_26'	[1,7,8]	76,19%	94,83%
'newff_27'	[1,7,9]	93,90%	94,95%	'newff_27'	[1,7,9]	80,95%	94,85%
'newff_28'	[1,8,9]	96,34%	94,97%	'newff_28'	[1,8,9]	85,71%	94,87%
'newff_29'	[2,3,4]	86,59%	94,80%	'newff_29'	[2,3,4]	83,33%	94,76%
'newff_30'	[2,3,5]	89,02%	94,89%	'newff_30'	[2,3,5]	85,71%	94,89%
'newff_31'	[2,3,6]	89,02%	94,85%	'newff_31'	[2,3,6]	73,81%	94,87%
'newff_32'	[2,3,7]	85,37%	94,75%	'newff_32'	[2,3,7]	78,57%	94,80%
'newff_33'	[2,3,8]	80,49%	94,63%	'newff_33'	[2,3,8]	71,43%	94,84%
'newff_34'	[2,3,9]	82,93%	94,82%	'newff_34'	[2,3,9]	78,57%	94,98%
'newff_35'	[2,4,5]	90,24%	94,86%	'newff_35'	[2,4,5]	83,33%	94,83%
'newff_36'	[2,4,6]	82,93%	93,76%	'newff_36'	[2,4,6]	54,76%	93,96%
'newff_37'	[2,4,7]	85,37%	94,99%	'newff_37'	[2,4,7]	76,19%	94,87%
'newff_38'	[2,4,8]	82,93%	94,89%	'newff_38'	[2,4,8]	66,67%	94,16%
'newff_39'	[2,4,9]	78,05%	94,97%	'newff_39'	[2,4,9]	71,43%	94,97%
'newff_40'	[2,5,6]	92,68%	94,92%	'newff_40'	[2,5,6]	90,48%	94,93%
'newff_41'	[2,5,7]	90,24%	94,86%	'newff_41'	[2,5,7]	85,71%	94,99%
'newff_42'	[2,5,8]	87,80%	92,81%	'newff_42'	[2,5,8]	85,71%	92,74%
'newff_43'	[2,5,9]	91,46%	93,99%	'newff_43'	[2,5,9]	90,48%	94,09%
'newff_44'	[2,6,7]	80,49%	94,85%	'newff_44'	[2,6,7]	71,43%	94,88%
'newff_45'	[2,6,8]	84,15%	94,86%	'newff_45'	[2,6,8]	69,05%	94,99%
'newff_46'	[2,6,9]	82,93%	94,95%	'newff_46'	[2,6,9]	73,81%	94,98%
'newff_47'	[2,7,8]	80,49%	94,99%	'newff_47'	[2,7,8]	69,05%	94,92%
'newff_48'	[2,7,9]	85,37%	94,91%	'newff_48'	[2,7,9]	85,71%	94,79%
'newff_49'	[2,8,9]	84,15%	94,85%	'newff_49'	[2,8,9]	83,33%	94,92%
'newff_50'	[3,4,5]	91,46%	94,74%	'newff_50'	[3,4,5]	90,48%	94,67%
'newff_51'	[3,4,6]	74,39%	93,79%	'newff_51'	[3,4,6]	66,67%	93,99%
'newff_52'	[3,4,7]	87,80%	94,69%	'newff_52'	[3,4,7]	64,29%	94,74%
'newff_53'	[3,4,8]	73,17%	99,52%	'newff_53'	[3,4,8]	61,90%	99,42%
'newff_54'	[3,4,9]	90,24%	94,98%	'newff_54'	[3,4,9]	85,71%	94,55%
'newff_55'	[3,5,6]	89,02%	94,81%	'newff_55'	[3,5,6]	80,95%	94,84%
'newff_56'	[3,5,7]	92,68%	94,83%	'newff_56'	[3,5,7]	88,10%	94,95%
'newff_57'	[3,5,8]	89,02%	94,98%	'newff_57'	[3,5,8]	85,71%	94,51%
'newff_58'	[3,5,9]	82,93%	94,59%	'newff_58'	[3,5,9]	78,57%	94,84%
'newff_59'	[3,6,7]	92,68%	94,62%	'newff_59'	[3,6,7]	85,71%	94,80%
'newff_60'	[3,6,8]	78,05%	99,11%	'newff_60'	[3,6,8]	57,14%	98,87%
'newff_61'	[3,6,9]	91,46%	94,77%	'newff_61'	[3,6,9]	92,86%	94,89%
'newff_62'	[3,7,8]	95,12%	94,83%	'newff_62'	[3,7,8]	90,48%	94,43%
'newff_63'	[3,7,9]	90,24%	94,77%	'newff_63'	[3,7,9]	78,57%	94,60%
'newff_64'	[3,8,9]	92,68%	94,98%	'newff_64'	[3,8,9]	95,24%	94,80%
'newff_65'	[4,5,6]	89,02%	94,64%	'newff_65'	[4,5,6]	92,86%	94,81%
'newff_66'	[4,5,7]	91,46%	98,46%	'newff_66'	[4,5,7]	92,86%	98,78%
'newff_67'	[4,5,8]	96,34%	94,95%	'newff_67'	[4,5,8]	83,33%	94,83%
'newff_68'	[4,5,9]	76,83%	99,59%	'newff_68'	[4,5,9]	73,81%	99,49%
'newff_69'	[4,6,7]	90,24%	94,95%	'newff_69'	[4,6,7]	83,33%	94,82%
'newff_70'	[4,6,8]	87,80%	94,89%	'newff_70'	[4,6,8]	80,95%	94,87%
'newff_71'	[4,6,9]	87,80%	94,69%	'newff_71'	[4,6,9]	85,71%	94,22%

'newff_72'	[4,7,8]	92,68%	94,90%		'newff_72'	[4,7,8]	76,19%	94,89%
'newff_73'	[4,7,9]	91,46%	94,77%		'newff_73'	[4,7,9]	80,95%	94,97%
'newff_74'	[4,8,9]	89,02%	94,88%		'newff_74'	[4,8,9]	90,48%	94,86%
'newff_75'	[5,6,7]	93,90%	94,96%		'newff_75'	[5,6,7]	88,10%	94,78%
'newff_76'	[5,6,8]	91,46%	94,27%		'newff_76'	[5,6,8]	88,10%	94,56%
'newff_77'	[5,6,9]	92,68%	94,18%		'newff_77'	[5,6,9]	80,95%	94,35%
'newff_78'	[5,7,8]	95,12%	94,85%		'newff_78'	[5,7,8]	90,48%	94,72%
'newff_79'	[5,7,9]	96,34%	94,70%		'newff_79'	[5,7,9]	80,95%	94,74%
'newff_80'	[5,8,9]	87,80%	94,90%		'newff_80'	[5,8,9]	78,57%	94,95%
'newff_81'	[6,7,8]	90,24%	94,94%		'newff_81'	[6,7,8]	83,33%	94,89%
'newff_82'	[6,7,9]	90,24%	94,73%		'newff_82'	[6,7,9]	92,86%	94,96%
'newff_83'	[6,8,9]	91,46%	94,99%		'newff_83'	[6,8,9]	92,86%	94,93%
'newff_84'	[7,8,9]	90,24%	94,98%		'newff_84'	[7,8,9]	78,57%	94,96%
Test 4 (4 features left out):								
Training Results:					Validation Results:			
Model ID	Features left out	T21	Normal		Model ID	Features left out	T21	Normal
'newff_1'	[1,2,3,4]	78,05%	98,52%		'newff_1'	[1,2,3,4]	66,67%	98,42%
'newff_2'	[1,2,3,5]	90,24%	94,67%		'newff_2'	[1,2,3,5]	85,71%	94,93%
'newff_3'	[1,2,3,6]	84,15%	94,92%		'newff_3'	[1,2,3,6]	76,19%	94,90%
'newff_4'	[1,2,3,7]	85,37%	94,94%		'newff_4'	[1,2,3,7]	83,33%	94,96%
'newff_5'	[1,2,3,8]	85,37%	94,75%		'newff_5'	[1,2,3,8]	73,81%	94,77%
'newff_6'	[1,2,3,9]	78,05%	94,96%		'newff_6'	[1,2,3,9]	64,29%	94,84%
'newff_7'	[1,2,4,5]	89,02%	94,67%		'newff_7'	[1,2,4,5]	78,57%	94,64%
'newff_8'	[1,2,4,6]	82,93%	94,84%		'newff_8'	[1,2,4,6]	78,57%	94,98%
'newff_9'	[1,2,4,7]	75,61%	94,78%		'newff_9'	[1,2,4,7]	73,81%	94,88%
'newff_10'	[1,2,4,8]	86,59%	94,79%		'newff_10'	[1,2,4,8]	61,90%	94,98%
'newff_11'	[1,2,4,9]	81,71%	94,19%		'newff_11'	[1,2,4,9]	78,57%	94,41%
'newff_12'	[1,2,5,6]	90,24%	94,89%		'newff_12'	[1,2,5,6]	83,33%	94,77%
'newff_13'	[1,2,5,7]	84,15%	94,73%		'newff_13'	[1,2,5,7]	80,95%	94,19%
'newff_14'	[1,2,5,8]	85,37%	94,91%		'newff_14'	[1,2,5,8]	80,95%	94,86%
'newff_15'	[1,2,5,9]	85,37%	94,98%		'newff_15'	[1,2,5,9]	83,33%	94,65%
'newff_16'	[1,2,6,7]	82,93%	94,90%		'newff_16'	[1,2,6,7]	66,67%	94,96%
'newff_17'	[1,2,6,8]	80,49%	94,76%		'newff_17'	[1,2,6,8]	69,05%	94,87%
'newff_18'	[1,2,6,9]	67,07%	97,71%		'newff_18'	[1,2,6,9]	54,76%	97,49%
'newff_19'	[1,2,7,8]	79,27%	94,80%		'newff_19'	[1,2,7,8]	69,05%	94,79%
'newff_20'	[1,2,7,9]	78,05%	94,92%		'newff_20'	[1,2,7,9]	69,05%	94,89%
'newff_21'	[1,2,8,9]	80,49%	93,98%		'newff_21'	[1,2,8,9]	73,81%	94,32%
'newff_22'	[1,3,4,5]	95,12%	94,65%		'newff_22'	[1,3,4,5]	90,48%	94,68%
'newff_23'	[1,3,4,6]	80,49%	98,21%		'newff_23'	[1,3,4,6]	71,43%	98,27%
'newff_24'	[1,3,4,7]	95,12%	94,93%		'newff_24'	[1,3,4,7]	92,86%	94,82%
'newff_25'	[1,3,4,8]	89,02%	94,65%		'newff_25'	[1,3,4,8]	83,33%	94,45%
'newff_26'	[1,3,4,9]	81,71%	94,96%		'newff_26'	[1,3,4,9]	73,81%	94,98%
'newff_27'	[1,3,5,6]	95,12%	94,79%		'newff_27'	[1,3,5,6]	85,71%	94,61%
'newff_28'	[1,3,5,7]	96,34%	94,67%		'newff_28'	[1,3,5,7]	92,86%	94,81%
'newff_29'	[1,3,5,8]	91,46%	94,83%		'newff_29'	[1,3,5,8]	85,71%	94,55%
'newff_30'	[1,3,5,9]	89,02%	94,75%		'newff_30'	[1,3,5,9]	90,48%	94,90%

'newff_31'	[1,3,6,7]	91,46%	94,85%	'newff_31'	[1,3,6,7]	80,95%	94,62%
'newff_32'	[1,3,6,8]	90,24%	94,67%	'newff_32'	[1,3,6,8]	90,48%	94,93%
'newff_33'	[1,3,6,9]	91,46%	94,72%	'newff_33'	[1,3,6,9]	85,71%	94,96%
'newff_34'	[1,3,7,8]	93,90%	94,83%	'newff_34'	[1,3,7,8]	90,48%	94,76%
'newff_35'	[1,3,7,9]	92,68%	94,61%	'newff_35'	[1,3,7,9]	95,24%	94,69%
'newff_36'	[1,3,8,9]	90,24%	94,46%	'newff_36'	[1,3,8,9]	88,10%	94,97%
'newff_37'	[1,4,5,6]	90,24%	94,86%	'newff_37'	[1,4,5,6]	85,71%	94,84%
'newff_38'	[1,4,5,7]	95,12%	94,83%	'newff_38'	[1,4,5,7]	92,86%	94,87%
'newff_39'	[1,4,5,8]	95,12%	93,10%	'newff_39'	[1,4,5,8]	83,33%	93,56%
'newff_40'	[1,4,5,9]	95,12%	94,90%	'newff_40'	[1,4,5,9]	85,71%	94,82%
'newff_41'	[1,4,6,7]	87,80%	94,88%	'newff_41'	[1,4,6,7]	83,33%	94,88%
'newff_42'	[1,4,6,8]	76,83%	99,04%	'newff_42'	[1,4,6,8]	59,52%	99,02%
'newff_43'	[1,4,6,9]	84,15%	94,87%	'newff_43'	[1,4,6,9]	71,43%	94,80%
'newff_44'	[1,4,7,8]	91,46%	94,82%	'newff_44'	[1,4,7,8]	83,33%	94,95%
'newff_45'	[1,4,7,9]	93,90%	94,97%	'newff_45'	[1,4,7,9]	97,62%	94,77%
'newff_46'	[1,4,8,9]	92,68%	94,82%	'newff_46'	[1,4,8,9]	73,81%	94,62%
'newff_47'	[1,5,6,7]	96,34%	94,60%	'newff_47'	[1,5,6,7]	90,48%	94,62%
'newff_48'	[1,5,6,8]	93,90%	94,84%	'newff_48'	[1,5,6,8]	85,71%	94,58%
'newff_49'	[1,5,6,9]	89,02%	94,73%	'newff_49'	[1,5,6,9]	95,24%	94,79%
'newff_50'	[1,5,7,8]	97,56%	94,78%	'newff_50'	[1,5,7,8]	92,86%	94,97%
'newff_51'	[1,5,7,9]	91,46%	94,95%	'newff_51'	[1,5,7,9]	90,48%	94,99%
'newff_52'	[1,5,8,9]	80,49%	99,40%	'newff_52'	[1,5,8,9]	73,81%	99,29%
'newff_53'	[1,6,7,8]	92,68%	94,74%	'newff_53'	[1,6,7,8]	80,95%	94,44%
'newff_54'	[1,6,7,9]	91,46%	94,85%	'newff_54'	[1,6,7,9]	88,10%	94,34%
'newff_55'	[1,6,8,9]	91,46%	94,95%	'newff_55'	[1,6,8,9]	83,33%	94,46%
'newff_56'	[1,7,8,9]	93,90%	94,88%	'newff_56'	[1,7,8,9]	85,71%	94,85%
'newff_57'	[2,3,4,5]	84,15%	94,73%	'newff_57'	[2,3,4,5]	85,71%	94,66%
'newff_58'	[2,3,4,6]	81,71%	94,75%	'newff_58'	[2,3,4,6]	66,67%	94,61%
'newff_59'	[2,3,4,7]	75,61%	94,94%	'newff_59'	[2,3,4,7]	64,29%	94,74%
'newff_60'	[2,3,4,8]	80,49%	94,97%	'newff_60'	[2,3,4,8]	73,81%	94,85%
'newff_61'	[2,3,4,9]	76,83%	94,72%	'newff_61'	[2,3,4,9]	76,19%	94,95%
'newff_62'	[2,3,5,6]	85,37%	94,73%	'newff_62'	[2,3,5,6]	76,19%	94,71%
'newff_63'	[2,3,5,7]	85,37%	94,33%	'newff_63'	[2,3,5,7]	71,43%	94,44%
'newff_64'	[2,3,5,8]	78,05%	94,88%	'newff_64'	[2,3,5,8]	73,81%	94,76%
'newff_65'	[2,3,5,9]	79,27%	94,12%	'newff_65'	[2,3,5,9]	85,71%	94,08%
'newff_66'	[2,3,6,7]	80,49%	94,48%	'newff_66'	[2,3,6,7]	71,43%	94,33%
'newff_67'	[2,3,6,8]	73,17%	94,91%	'newff_67'	[2,3,6,8]	54,76%	94,76%
'newff_68'	[2,3,6,9]	52,44%	99,28%	'newff_68'	[2,3,6,9]	42,86%	99,13%
'newff_69'	[2,3,7,8]	76,83%	94,74%	'newff_69'	[2,3,7,8]	47,62%	94,87%
'newff_70'	[2,3,7,9]	82,93%	94,50%	'newff_70'	[2,3,7,9]	83,33%	94,45%
'newff_71'	[2,3,8,9]	75,61%	94,63%	'newff_71'	[2,3,8,9]	71,43%	94,63%
'newff_72'	[2,4,5,6]	84,15%	94,98%	'newff_72'	[2,4,5,6]	76,19%	94,94%
'newff_73'	[2,4,5,7]	85,37%	94,89%	'newff_73'	[2,4,5,7]	83,33%	94,95%
'newff_74'	[2,4,5,8]	82,93%	94,93%	'newff_74'	[2,4,5,8]	69,05%	94,97%
'newff_75'	[2,4,5,9]	82,93%	94,55%	'newff_75'	[2,4,5,9]	69,05%	94,91%
'newff_76'	[2,4,6,7]	71,95%	94,96%	'newff_76'	[2,4,6,7]	64,29%	94,88%
'newff_77'	[2,4,6,8]	74,39%	94,93%	'newff_77'	[2,4,6,8]	54,76%	94,77%

'newff_78'	[2,4,6,9]	74,39%	94,73%		'newff_78'	[2,4,6,9]	71,43%	94,79%
'newff_79'	[2,4,7,8]	75,61%	94,33%		'newff_79'	[2,4,7,8]	54,76%	94,61%
'newff_80'	[2,4,7,9]	76,83%	94,84%		'newff_80'	[2,4,7,9]	85,71%	94,96%
'newff_81'	[2,4,8,9]	76,83%	94,71%		'newff_81'	[2,4,8,9]	73,81%	94,95%
'newff_82'	[2,5,6,7]	84,15%	95,00%		'newff_82'	[2,5,6,7]	76,19%	94,94%
'newff_83'	[2,5,6,8]	84,15%	94,87%		'newff_83'	[2,5,6,8]	76,19%	94,74%
'newff_84'	[2,5,6,9]	85,37%	94,92%		'newff_84'	[2,5,6,9]	78,57%	94,90%
'newff_85'	[2,5,7,8]	80,49%	94,12%		'newff_85'	[2,5,7,8]	69,05%	94,57%
'newff_86'	[2,5,7,9]	89,02%	94,84%		'newff_86'	[2,5,7,9]	90,48%	94,91%
'newff_87'	[2,5,8,9]	81,71%	94,70%		'newff_87'	[2,5,8,9]	78,57%	94,58%
'newff_88'	[2,6,7,8]	69,51%	94,95%		'newff_88'	[2,6,7,8]	61,90%	94,84%
'newff_89'	[2,6,7,9]	78,05%	94,92%		'newff_89'	[2,6,7,9]	76,19%	94,96%
'newff_90'	[2,6,8,9]	78,05%	94,96%		'newff_90'	[2,6,8,9]	71,43%	94,88%
'newff_91'	[2,7,8,9]	82,93%	94,97%		'newff_91'	[2,7,8,9]	78,57%	94,81%
'newff_92'	[3,4,5,6]	87,80%	93,47%		'newff_92'	[3,4,5,6]	88,10%	93,72%
'newff_93'	[3,4,5,7]	95,12%	94,62%		'newff_93'	[3,4,5,7]	90,48%	94,67%
'newff_94'	[3,4,5,8]	87,80%	94,85%		'newff_94'	[3,4,5,8]	85,71%	94,70%
'newff_95'	[3,4,5,9]	82,93%	94,12%		'newff_95'	[3,4,5,9]	71,43%	94,96%
'newff_96'	[3,4,6,7]	80,49%	94,91%		'newff_96'	[3,4,6,7]	88,10%	94,86%
'newff_97'	[3,4,6,8]	86,59%	94,42%		'newff_97'	[3,4,6,8]	80,95%	94,84%
'newff_98'	[3,4,6,9]	92,68%	47,68%		'newff_98'	[3,4,6,9]	76,19%	48,12%
'newff_99'	[3,4,7,8]	85,37%	94,93%		'newff_99'	[3,4,7,8]	90,48%	94,93%
'newff_100'	[3,4,7,9]	87,80%	93,58%		'newff_100'	[3,4,7,9]	90,48%	94,02%
'newff_101'	[3,4,8,9]	89,02%	94,51%		'newff_101'	[3,4,8,9]	88,10%	94,90%
'newff_102'	[3,5,6,7]	91,46%	94,86%		'newff_102'	[3,5,6,7]	95,24%	94,92%
'newff_103'	[3,5,6,8]	84,15%	94,92%		'newff_103'	[3,5,6,8]	78,57%	94,74%
'newff_104'	[3,5,6,9]	89,02%	94,06%		'newff_104'	[3,5,6,9]	83,33%	94,07%
'newff_105'	[3,5,7,8]	90,24%	94,96%		'newff_105'	[3,5,7,8]	88,10%	94,85%
'newff_106'	[3,5,7,9]	87,80%	94,96%		'newff_106'	[3,5,7,9]	88,10%	94,87%
'newff_107'	[3,5,8,9]	85,37%	94,93%		'newff_107'	[3,5,8,9]	90,48%	94,95%
'newff_108'	[3,6,7,8]	81,71%	94,81%		'newff_108'	[3,6,7,8]	73,81%	94,85%
'newff_109'	[3,6,7,9]	89,02%	94,52%		'newff_109'	[3,6,7,9]	88,10%	94,54%
'newff_110'	[3,6,8,9]	86,59%	95,00%		'newff_110'	[3,6,8,9]	85,71%	94,64%
'newff_111'	[3,7,8,9]	87,80%	94,80%		'newff_111'	[3,7,8,9]	85,71%	94,92%
'newff_112'	[4,5,6,7]	87,80%	94,96%		'newff_112'	[4,5,6,7]	90,48%	94,61%
'newff_113'	[4,5,6,8]	86,59%	94,99%		'newff_113'	[4,5,6,8]	66,67%	94,94%
'newff_114'	[4,5,6,9]	84,15%	94,51%		'newff_114'	[4,5,6,9]	78,57%	94,81%
'newff_115'	[4,5,7,8]	90,24%	94,90%		'newff_115'	[4,5,7,8]	83,33%	94,69%
'newff_116'	[4,5,7,9]	84,15%	94,95%		'newff_116'	[4,5,7,9]	78,57%	94,80%
'newff_117'	[4,5,8,9]	89,02%	94,53%		'newff_117'	[4,5,8,9]	88,10%	94,89%
'newff_118'	[4,6,7,8]	84,15%	94,84%		'newff_118'	[4,6,7,8]	71,43%	94,72%
'newff_119'	[4,6,7,9]	84,15%	94,97%		'newff_119'	[4,6,7,9]	83,33%	94,69%
'newff_120'	[4,6,8,9]	69,51%	99,16%		'newff_120'	[4,6,8,9]	71,43%	99,11%
'newff_121'	[4,7,8,9]	95,12%	83,68%		'newff_121'	[4,7,8,9]	78,57%	83,89%
'newff_122'	[5,6,7,8]	92,68%	94,89%		'newff_122'	[5,6,7,8]	83,33%	94,89%
'newff_123'	[5,6,7,9]	91,46%	94,91%		'newff_123'	[5,6,7,9]	92,86%	94,81%
'newff_124'	[5,6,8,9]	81,71%	94,71%		'newff_124'	[5,6,8,9]	83,33%	94,69%

'newff_125'	[5,7,8,9]	84,15%	94,77%		'newff_125'	[5,7,8,9]	78,57%	93,32%
'newff_126'	[6,7,8,9]	90,24%	94,99%		'newff_126'	[6,7,8,9]	76,19%	94,92%
Test 5 (5 features left out):								
Training Results:					Validation Results:			
Model ID	Features left out	T21	Normal		Model ID	Features left out	T21	Normal
'newff_1'	[1,2,3,4,5]	86,59%	94,70%		'newff_1'	[1,2,3,4,5]	83,33%	94,99%
'newff_2'	[1,2,3,4,6]	79,27%	94,93%		'newff_2'	[1,2,3,4,6]	61,90%	94,88%
'newff_3'	[1,2,3,4,7]	79,27%	94,67%		'newff_3'	[1,2,3,4,7]	69,05%	94,80%
'newff_4'	[1,2,3,4,8]	76,83%	95,00%		'newff_4'	[1,2,3,4,8]	71,43%	94,96%
'newff_5'	[1,2,3,4,9]	78,05%	94,98%		'newff_5'	[1,2,3,4,9]	73,81%	94,86%
'newff_6'	[1,2,3,5,6]	81,71%	94,95%		'newff_6'	[1,2,3,5,6]	76,19%	94,77%
'newff_7'	[1,2,3,5,7]	80,49%	94,88%		'newff_7'	[1,2,3,5,7]	85,71%	94,96%
'newff_8'	[1,2,3,5,8]	78,05%	94,97%		'newff_8'	[1,2,3,5,8]	76,19%	94,95%
'newff_9'	[1,2,3,5,9]	81,71%	94,83%		'newff_9'	[1,2,3,5,9]	73,81%	94,89%
'newff_10'	[1,2,3,6,7]	78,05%	94,99%		'newff_10'	[1,2,3,6,7]	64,29%	94,99%
'newff_11'	[1,2,3,6,8]	74,39%	94,89%		'newff_11'	[1,2,3,6,8]	64,29%	94,95%
'newff_12'	[1,2,3,6,9]	75,61%	94,88%		'newff_12'	[1,2,3,6,9]	80,95%	94,83%
'newff_13'	[1,2,3,7,8]	75,61%	94,86%		'newff_13'	[1,2,3,7,8]	61,90%	94,88%
'newff_14'	[1,2,3,7,9]	78,05%	94,94%		'newff_14'	[1,2,3,7,9]	76,19%	94,94%
'newff_15'	[1,2,3,8,9]	76,83%	95,00%		'newff_15'	[1,2,3,8,9]	71,43%	94,73%
'newff_16'	[1,2,4,5,6]	84,15%	94,99%		'newff_16'	[1,2,4,5,6]	76,19%	94,90%
'newff_17'	[1,2,4,5,7]	80,49%	94,60%		'newff_17'	[1,2,4,5,7]	85,71%	94,51%
'newff_18'	[1,2,4,5,8]	84,15%	94,81%		'newff_18'	[1,2,4,5,8]	71,43%	94,71%
'newff_19'	[1,2,4,5,9]	78,05%	94,98%		'newff_19'	[1,2,4,5,9]	73,81%	94,88%
'newff_20'	[1,2,4,6,7]	67,07%	94,91%		'newff_20'	[1,2,4,6,7]	61,90%	94,86%
'newff_21'	[1,2,4,6,8]	74,39%	94,99%		'newff_21'	[1,2,4,6,8]	59,52%	94,96%
'newff_22'	[1,2,4,6,9]	80,49%	95,00%		'newff_22'	[1,2,4,6,9]	73,81%	94,88%
'newff_23'	[1,2,4,7,8]	71,95%	95,00%		'newff_23'	[1,2,4,7,8]	73,81%	94,98%
'newff_24'	[1,2,4,7,9]	74,39%	94,72%		'newff_24'	[1,2,4,7,9]	73,81%	94,86%
'newff_25'	[1,2,4,8,9]	76,83%	94,90%		'newff_25'	[1,2,4,8,9]	73,81%	94,89%
'newff_26'	[1,2,5,6,7]	80,49%	94,88%		'newff_26'	[1,2,5,6,7]	69,05%	94,98%
'newff_27'	[1,2,5,6,8]	81,71%	94,81%		'newff_27'	[1,2,5,6,8]	78,57%	94,86%
'newff_28'	[1,2,5,6,9]	79,27%	94,85%		'newff_28'	[1,2,5,6,9]	73,81%	94,66%
'newff_29'	[1,2,5,7,8]	79,27%	94,72%		'newff_29'	[1,2,5,7,8]	78,57%	94,77%
'newff_30'	[1,2,5,7,9]	84,15%	94,97%		'newff_30'	[1,2,5,7,9]	78,57%	94,87%
'newff_31'	[1,2,5,8,9]	82,93%	94,85%		'newff_31'	[1,2,5,8,9]	76,19%	94,95%
'newff_32'	[1,2,6,7,8]	70,73%	94,84%		'newff_32'	[1,2,6,7,8]	54,76%	94,72%
'newff_33'	[1,2,6,7,9]	74,39%	94,72%		'newff_33'	[1,2,6,7,9]	64,29%	94,85%
'newff_34'	[1,2,6,8,9]	74,39%	94,97%		'newff_34'	[1,2,6,8,9]	66,67%	94,98%
'newff_35'	[1,2,7,8,9]	75,61%	94,79%		'newff_35'	[1,2,7,8,9]	57,14%	94,85%
'newff_36'	[1,3,4,5,6]	91,46%	94,76%		'newff_36'	[1,3,4,5,6]	88,10%	94,66%
'newff_37'	[1,3,4,5,7]	92,68%	94,78%		'newff_37'	[1,3,4,5,7]	90,48%	94,78%
'newff_38'	[1,3,4,5,8]	89,02%	94,73%		'newff_38'	[1,3,4,5,8]	85,71%	94,79%
'newff_39'	[1,3,4,5,9]	87,80%	94,89%		'newff_39'	[1,3,4,5,9]	88,10%	94,43%
'newff_40'	[1,3,4,6,7]	87,80%	94,91%		'newff_40'	[1,3,4,6,7]	95,24%	94,84%
'newff_41'	[1,3,4,6,8]	87,80%	88,05%		'newff_41'	[1,3,4,6,8]	78,57%	87,61%

'newff_42'	[1,3,4,6,9]	89,02%	93,54%		'newff_42'	[1,3,4,6,9]	83,33%	93,92%
'newff_43'	[1,3,4,7,8]	90,24%	94,83%		'newff_43'	[1,3,4,7,8]	88,10%	94,73%
'newff_44'	[1,3,4,7,9]	89,02%	94,88%		'newff_44'	[1,3,4,7,9]	92,86%	94,90%
'newff_45'	[1,3,4,8,9]	87,80%	94,96%		'newff_45'	[1,3,4,8,9]	83,33%	94,91%
'newff_46'	[1,3,5,6,7]	91,46%	94,80%		'newff_46'	[1,3,5,6,7]	90,48%	94,74%
'newff_47'	[1,3,5,6,8]	87,80%	94,58%		'newff_47'	[1,3,5,6,8]	78,57%	94,87%
'newff_48'	[1,3,5,6,9]	73,17%	99,52%		'newff_48'	[1,3,5,6,9]	61,90%	99,42%
'newff_49'	[1,3,5,7,8]	91,46%	94,96%		'newff_49'	[1,3,5,7,8]	88,10%	94,80%
'newff_50'	[1,3,5,7,9]	91,46%	94,96%		'newff_50'	[1,3,5,7,9]	95,24%	94,56%
'newff_51'	[1,3,5,8,9]	91,46%	94,87%		'newff_51'	[1,3,5,8,9]	90,48%	94,84%
'newff_52'	[1,3,6,7,8]	85,37%	94,96%		'newff_52'	[1,3,6,7,8]	80,95%	94,87%
'newff_53'	[1,3,6,7,9]	86,59%	94,64%		'newff_53'	[1,3,6,7,9]	92,86%	94,54%
'newff_54'	[1,3,6,8,9]	85,37%	94,89%		'newff_54'	[1,3,6,8,9]	80,95%	94,57%
'newff_55'	[1,3,7,8,9]	91,46%	94,94%		'newff_55'	[1,3,7,8,9]	90,48%	94,91%
'newff_56'	[1,4,5,6,7]	89,02%	94,94%		'newff_56'	[1,4,5,6,7]	85,71%	94,84%
'newff_57'	[1,4,5,6,8]	86,59%	94,70%		'newff_57'	[1,4,5,6,8]	76,19%	94,77%
'newff_58'	[1,4,5,6,9]	86,59%	94,94%		'newff_58'	[1,4,5,6,9]	88,10%	94,87%
'newff_59'	[1,4,5,7,8]	97,56%	94,88%		'newff_59'	[1,4,5,7,8]	92,86%	94,85%
'newff_60'	[1,4,5,7,9]	91,46%	94,81%		'newff_60'	[1,4,5,7,9]	85,71%	94,82%
'newff_61'	[1,4,5,8,9]	90,24%	94,97%		'newff_61'	[1,4,5,8,9]	90,48%	94,84%
'newff_62'	[1,4,6,7,8]	82,93%	94,85%		'newff_62'	[1,4,6,7,8]	76,19%	94,89%
'newff_63'	[1,4,6,7,9]	84,15%	94,87%		'newff_63'	[1,4,6,7,9]	80,95%	94,66%
'newff_64'	[1,4,6,8,9]	86,59%	94,85%		'newff_64'	[1,4,6,8,9]	78,57%	94,76%
'newff_65'	[1,4,7,8,9]	85,37%	95,00%		'newff_65'	[1,4,7,8,9]	85,71%	94,78%
'newff_66'	[1,5,6,7,8]	89,02%	94,81%		'newff_66'	[1,5,6,7,8]	85,71%	94,95%
'newff_67'	[1,5,6,7,9]	89,02%	94,97%		'newff_67'	[1,5,6,7,9]	90,48%	94,61%
'newff_68'	[1,5,6,8,9]	92,68%	94,80%		'newff_68'	[1,5,6,8,9]	90,48%	94,89%
'newff_69'	[1,5,7,8,9]	86,59%	94,91%		'newff_69'	[1,5,7,8,9]	83,33%	94,70%
'newff_70'	[1,6,7,8,9]	91,46%	94,91%		'newff_70'	[1,6,7,8,9]	90,48%	94,95%
'newff_71'	[2,3,4,5,6]	82,93%	94,93%		'newff_71'	[2,3,4,5,6]	73,81%	94,91%
'newff_72'	[2,3,4,5,7]	78,05%	94,87%		'newff_72'	[2,3,4,5,7]	80,95%	94,75%
'newff_73'	[2,3,4,5,8]	74,39%	94,87%		'newff_73'	[2,3,4,5,8]	61,90%	94,76%
'newff_74'	[2,3,4,5,9]	82,93%	94,39%		'newff_74'	[2,3,4,5,9]	80,95%	94,79%
'newff_75'	[2,3,4,6,7]	71,95%	94,94%		'newff_75'	[2,3,4,6,7]	57,14%	94,87%
'newff_76'	[2,3,4,6,8]	70,73%	94,94%		'newff_76'	[2,3,4,6,8]	57,14%	94,90%
'newff_77'	[2,3,4,6,9]	75,61%	94,71%		'newff_77'	[2,3,4,6,9]	69,05%	94,92%
'newff_78'	[2,3,4,7,8]	65,85%	94,98%		'newff_78'	[2,3,4,7,8]	42,86%	94,86%
'newff_79'	[2,3,4,7,9]	75,61%	94,89%		'newff_79'	[2,3,4,7,9]	76,19%	94,92%
'newff_80'	[2,3,4,8,9]	80,49%	87,29%		'newff_80'	[2,3,4,8,9]	78,57%	87,47%
'newff_81'	[2,3,5,6,7]	79,27%	94,70%		'newff_81'	[2,3,5,6,7]	61,90%	94,78%
'newff_82'	[2,3,5,6,8]	74,39%	94,83%		'newff_82'	[2,3,5,6,8]	66,67%	94,94%
'newff_83'	[2,3,5,6,9]	75,61%	94,89%		'newff_83'	[2,3,5,6,9]	80,95%	94,77%
'newff_84'	[2,3,5,7,8]	78,05%	94,96%		'newff_84'	[2,3,5,7,8]	64,29%	94,96%
'newff_85'	[2,3,5,7,9]	90,24%	88,94%		'newff_85'	[2,3,5,7,9]	76,19%	89,58%
'newff_86'	[2,3,5,8,9]	73,17%	94,79%		'newff_86'	[2,3,5,8,9]	76,19%	94,12%
'newff_87'	[2,3,6,7,8]	60,98%	94,98%		'newff_87'	[2,3,6,7,8]	40,48%	94,90%
'newff_88'	[2,3,6,7,9]	74,39%	94,90%		'newff_88'	[2,3,6,7,9]	59,52%	94,88%

'newff_89'	[2,3,6,8,9]	57,32%	94,56%		'newff_89'	[2,3,6,8,9]	57,14%	94,87%
'newff_90'	[2,3,7,8,9]	68,29%	94,89%		'newff_90'	[2,3,7,8,9]	66,67%	94,87%
'newff_91'	[2,4,5,6,7]	73,17%	94,80%		'newff_91'	[2,4,5,6,7]	66,67%	94,91%
'newff_92'	[2,4,5,6,8]	68,29%	94,90%		'newff_92'	[2,4,5,6,8]	54,76%	94,92%
'newff_93'	[2,4,5,6,9]	75,61%	94,88%		'newff_93'	[2,4,5,6,9]	80,95%	94,89%
'newff_94'	[2,4,5,7,8]	69,51%	94,69%		'newff_94'	[2,4,5,7,8]	80,95%	94,83%
'newff_95'	[2,4,5,7,9]	79,27%	94,83%		'newff_95'	[2,4,5,7,9]	78,57%	94,82%
'newff_96'	[2,4,5,8,9]	82,93%	94,73%		'newff_96'	[2,4,5,8,9]	76,19%	94,85%
'newff_97'	[2,4,6,7,8]	58,54%	94,77%		'newff_97'	[2,4,6,7,8]	38,10%	94,83%
'newff_98'	[2,4,6,7,9]	64,63%	94,82%		'newff_98'	[2,4,6,7,9]	64,29%	94,77%
'newff_99'	[2,4,6,8,9]	65,85%	94,95%		'newff_99'	[2,4,6,8,9]	54,76%	94,90%
'newff_100'	[2,4,7,8,9]	70,73%	94,88%		'newff_100'	[2,4,7,8,9]	69,05%	94,88%
'newff_101'	[2,5,6,7,8]	68,29%	94,85%		'newff_101'	[2,5,6,7,8]	71,43%	94,95%
'newff_102'	[2,5,6,7,9]	63,41%	94,94%		'newff_102'	[2,5,6,7,9]	73,81%	94,92%
'newff_103'	[2,5,6,8,9]	74,39%	94,91%		'newff_103'	[2,5,6,8,9]	71,43%	94,93%
'newff_104'	[2,5,7,8,9]	74,39%	94,68%		'newff_104'	[2,5,7,8,9]	73,81%	94,79%
'newff_105'	[2,6,7,8,9]	62,20%	94,83%		'newff_105'	[2,6,7,8,9]	54,76%	94,94%
'newff_106'	[3,4,5,6,7]	89,02%	94,94%		'newff_106'	[3,4,5,6,7]	85,71%	94,79%
'newff_107'	[3,4,5,6,8]	80,49%	94,85%		'newff_107'	[3,4,5,6,8]	83,33%	94,89%
'newff_108'	[3,4,5,6,9]	85,37%	94,89%		'newff_108'	[3,4,5,6,9]	78,57%	94,74%
'newff_109'	[3,4,5,7,8]	84,15%	94,95%		'newff_109'	[3,4,5,7,8]	80,95%	94,99%
'newff_110'	[3,4,5,7,9]	87,80%	94,94%		'newff_110'	[3,4,5,7,9]	88,10%	94,84%
'newff_111'	[3,4,5,8,9]	89,02%	94,69%		'newff_111'	[3,4,5,8,9]	90,48%	94,77%
'newff_112'	[3,4,6,7,8]	78,05%	94,96%		'newff_112'	[3,4,6,7,8]	73,81%	94,88%
'newff_113'	[3,4,6,7,9]	87,80%	94,81%		'newff_113'	[3,4,6,7,9]	88,10%	94,90%
'newff_114'	[3,4,6,8,9]	79,27%	93,92%		'newff_114'	[3,4,6,8,9]	73,81%	94,03%
'newff_115'	[3,4,7,8,9]	81,71%	94,90%		'newff_115'	[3,4,7,8,9]	83,33%	94,87%
'newff_116'	[3,5,6,7,8]	81,71%	94,97%		'newff_116'	[3,5,6,7,8]	80,95%	94,85%
'newff_117'	[3,5,6,7,9]	89,02%	94,71%		'newff_117'	[3,5,6,7,9]	95,24%	94,99%
'newff_118'	[3,5,6,8,9]	82,93%	94,51%		'newff_118'	[3,5,6,8,9]	85,71%	94,64%
'newff_119'	[3,5,7,8,9]	87,80%	94,99%		'newff_119'	[3,5,7,8,9]	88,10%	94,83%
'newff_120'	[3,6,7,8,9]	82,93%	94,84%		'newff_120'	[3,6,7,8,9]	76,19%	94,91%
'newff_121'	[4,5,6,7,8]	82,93%	94,86%		'newff_121'	[4,5,6,7,8]	83,33%	94,83%
'newff_122'	[4,5,6,7,9]	86,59%	94,90%		'newff_122'	[4,5,6,7,9]	95,24%	94,91%
'newff_123'	[4,5,6,8,9]	89,02%	94,70%		'newff_123'	[4,5,6,8,9]	85,71%	94,72%
'newff_124'	[4,5,7,8,9]	90,24%	94,80%		'newff_124'	[4,5,7,8,9]	88,10%	94,73%
'newff_125'	[4,6,7,8,9]	84,15%	94,80%		'newff_125'	[4,6,7,8,9]	78,57%	94,69%
'newff_126'	[5,6,7,8,9]	89,02%	94,89%		'newff_126'	[5,6,7,8,9]	88,10%	94,84%
Test 6 (6 features left out):								
Training Results:					Validation Results:			
Model ID	Features left out	T21	Normal		Model ID	Features left out	T21	Normal
'newff_1'	[1,2,3,4,5,6]	79,27%	94,95%		'newff_1'	[1,2,3,4,5,6]	71,43%	94,89%
'newff_2'	[1,2,3,4,5,7]	76,83%	94,77%		'newff_2'	[1,2,3,4,5,7]	76,19%	94,61%
'newff_3'	[1,2,3,4,5,8]	79,27%	94,85%		'newff_3'	[1,2,3,4,5,8]	73,81%	94,97%
'newff_4'	[1,2,3,4,5,9]	79,27%	94,80%		'newff_4'	[1,2,3,4,5,9]	80,95%	94,97%
'newff_5'	[1,2,3,4,6,7]	75,61%	94,84%		'newff_5'	[1,2,3,4,6,7]	59,52%	94,79%

'newff_6'	[1,2,3,4,6,8]	64,63%	94,87%	'newff_6'	[1,2,3,4,6,8]	52,38%	94,82%
'newff_7'	[1,2,3,4,6,9]	78,05%	94,95%	'newff_7'	[1,2,3,4,6,9]	76,19%	94,94%
'newff_8'	[1,2,3,4,7,8]	69,51%	94,93%	'newff_8'	[1,2,3,4,7,8]	69,05%	94,81%
'newff_9'	[1,2,3,4,7,9]	70,73%	94,82%	'newff_9'	[1,2,3,4,7,9]	73,81%	94,66%
'newff_10'	[1,2,3,4,8,9]	73,17%	94,83%	'newff_10'	[1,2,3,4,8,9]	76,19%	94,95%
'newff_11'	[1,2,3,5,6,7]	75,61%	94,74%	'newff_11'	[1,2,3,5,6,7]	61,90%	94,92%
'newff_12'	[1,2,3,5,6,8]	74,39%	94,99%	'newff_12'	[1,2,3,5,6,8]	66,67%	94,99%
'newff_13'	[1,2,3,5,6,9]	79,27%	94,74%	'newff_13'	[1,2,3,5,6,9]	76,19%	94,70%
'newff_14'	[1,2,3,5,7,8]	74,39%	94,95%	'newff_14'	[1,2,3,5,7,8]	69,05%	94,96%
'newff_15'	[1,2,3,5,7,9]	71,95%	94,93%	'newff_15'	[1,2,3,5,7,9]	85,71%	94,93%
'newff_16'	[1,2,3,5,8,9]	76,83%	94,95%	'newff_16'	[1,2,3,5,8,9]	71,43%	94,64%
'newff_17'	[1,2,3,6,7,8]	56,10%	94,80%	'newff_17'	[1,2,3,6,7,8]	38,10%	94,87%
'newff_18'	[1,2,3,6,7,9]	78,05%	94,84%	'newff_18'	[1,2,3,6,7,9]	66,67%	94,92%
'newff_19'	[1,2,3,6,8,9]	58,54%	94,49%	'newff_19'	[1,2,3,6,8,9]	57,14%	94,41%
'newff_20'	[1,2,3,7,8,9]	68,29%	94,95%	'newff_20'	[1,2,3,7,8,9]	64,29%	94,88%
'newff_21'	[1,2,4,5,6,7]	73,17%	94,78%	'newff_21'	[1,2,4,5,6,7]	71,43%	94,83%
'newff_22'	[1,2,4,5,6,8]	74,39%	94,93%	'newff_22'	[1,2,4,5,6,8]	61,90%	94,92%
'newff_23'	[1,2,4,5,6,9]	73,17%	94,95%	'newff_23'	[1,2,4,5,6,9]	76,19%	94,83%
'newff_24'	[1,2,4,5,7,8]	71,95%	94,93%	'newff_24'	[1,2,4,5,7,8]	64,29%	94,96%
'newff_25'	[1,2,4,5,7,9]	71,95%	94,98%	'newff_25'	[1,2,4,5,7,9]	76,19%	94,95%
'newff_26'	[1,2,4,5,8,9]	75,61%	94,86%	'newff_26'	[1,2,4,5,8,9]	76,19%	94,96%
'newff_27'	[1,2,4,6,7,8]	53,66%	94,84%	'newff_27'	[1,2,4,6,7,8]	30,95%	94,87%
'newff_28'	[1,2,4,6,7,9]	68,29%	94,98%	'newff_28'	[1,2,4,6,7,9]	64,29%	94,92%
'newff_29'	[1,2,4,6,8,9]	62,20%	94,99%	'newff_29'	[1,2,4,6,8,9]	52,38%	94,89%
'newff_30'	[1,2,4,7,8,9]	63,41%	94,87%	'newff_30'	[1,2,4,7,8,9]	69,05%	94,99%
'newff_31'	[1,2,5,6,7,8]	65,85%	94,92%	'newff_31'	[1,2,5,6,7,8]	59,52%	94,86%
'newff_32'	[1,2,5,6,7,9]	74,39%	94,88%	'newff_32'	[1,2,5,6,7,9]	71,43%	94,94%
'newff_33'	[1,2,5,6,8,9]	74,39%	94,91%	'newff_33'	[1,2,5,6,8,9]	66,67%	94,78%
'newff_34'	[1,2,5,7,8,9]	74,39%	94,98%	'newff_34'	[1,2,5,7,8,9]	57,14%	94,99%
'newff_35'	[1,2,6,7,8,9]	54,88%	94,90%	'newff_35'	[1,2,6,7,8,9]	47,62%	94,72%
'newff_36'	[1,3,4,5,6,7]	84,15%	94,74%	'newff_36'	[1,3,4,5,6,7]	83,33%	94,76%
'newff_37'	[1,3,4,5,6,8]	85,37%	94,78%	'newff_37'	[1,3,4,5,6,8]	78,57%	94,91%
'newff_38'	[1,3,4,5,6,9]	89,02%	94,75%	'newff_38'	[1,3,4,5,6,9]	92,86%	94,94%
'newff_39'	[1,3,4,5,7,8]	87,80%	94,50%	'newff_39'	[1,3,4,5,7,8]	85,71%	94,99%
'newff_40'	[1,3,4,5,7,9]	89,02%	95,00%	'newff_40'	[1,3,4,5,7,9]	92,86%	94,97%
'newff_41'	[1,3,4,5,8,9]	89,02%	94,81%	'newff_41'	[1,3,4,5,8,9]	90,48%	94,93%
'newff_42'	[1,3,4,6,7,8]	84,15%	94,97%	'newff_42'	[1,3,4,6,7,8]	80,95%	94,74%
'newff_43'	[1,3,4,6,7,9]	86,59%	94,67%	'newff_43'	[1,3,4,6,7,9]	85,71%	94,82%
'newff_44'	[1,3,4,6,8,9]	82,93%	94,69%	'newff_44'	[1,3,4,6,8,9]	78,57%	94,89%
'newff_45'	[1,3,4,7,8,9]	86,59%	94,71%	'newff_45'	[1,3,4,7,8,9]	85,71%	94,97%
'newff_46'	[1,3,5,6,7,8]	80,49%	94,83%	'newff_46'	[1,3,5,6,7,8]	85,71%	94,75%
'newff_47'	[1,3,5,6,7,9]	91,46%	95,00%	'newff_47'	[1,3,5,6,7,9]	90,48%	94,93%
'newff_48'	[1,3,5,6,8,9]	84,15%	94,84%	'newff_48'	[1,3,5,6,8,9]	85,71%	94,86%
'newff_49'	[1,3,5,7,8,9]	89,02%	94,90%	'newff_49'	[1,3,5,7,8,9]	90,48%	94,87%
'newff_50'	[1,3,6,7,8,9]	84,15%	94,82%	'newff_50'	[1,3,6,7,8,9]	80,95%	94,75%
'newff_51'	[1,4,5,6,7,8]	85,37%	94,87%	'newff_51'	[1,4,5,6,7,8]	80,95%	94,53%
'newff_52'	[1,4,5,6,7,9]	85,37%	94,92%	'newff_52'	[1,4,5,6,7,9]	92,86%	94,93%

'newff_53'	[1,4,5,6,8,9]	87,80%	94,95%	'newff_53'	[1,4,5,6,8,9]	76,19%	94,30%
'newff_54'	[1,4,5,7,8,9]	90,24%	94,85%	'newff_54'	[1,4,5,7,8,9]	88,10%	94,83%
'newff_55'	[1,4,6,7,8,9]	82,93%	94,76%	'newff_55'	[1,4,6,7,8,9]	78,57%	94,83%
'newff_56'	[1,5,6,7,8,9]	90,24%	94,98%	'newff_56'	[1,5,6,7,8,9]	88,10%	94,99%
'newff_57'	[2,3,4,5,6,7]	74,39%	94,99%	'newff_57'	[2,3,4,5,6,7]	57,14%	94,97%
'newff_58'	[2,3,4,5,6,8]	62,20%	94,80%	'newff_58'	[2,3,4,5,6,8]	57,14%	94,68%
'newff_59'	[2,3,4,5,6,9]	74,39%	94,97%	'newff_59'	[2,3,4,5,6,9]	80,95%	94,94%
'newff_60'	[2,3,4,5,7,8]	67,07%	94,86%	'newff_60'	[2,3,4,5,7,8]	69,05%	94,26%
'newff_61'	[2,3,4,5,7,9]	63,41%	94,94%	'newff_61'	[2,3,4,5,7,9]	59,52%	94,97%
'newff_62'	[2,3,4,5,8,9]	75,61%	94,97%	'newff_62'	[2,3,4,5,8,9]	76,19%	94,32%
'newff_63'	[2,3,4,6,7,8]	41,46%	94,86%	'newff_63'	[2,3,4,6,7,8]	23,81%	94,83%
'newff_64'	[2,3,4,6,7,9]	64,63%	94,88%	'newff_64'	[2,3,4,6,7,9]	52,38%	94,99%
'newff_65'	[2,3,4,6,8,9]	54,88%	94,52%	'newff_65'	[2,3,4,6,8,9]	50,00%	94,39%
'newff_66'	[2,3,4,7,8,9]	59,76%	94,05%	'newff_66'	[2,3,4,7,8,9]	64,29%	94,16%
'newff_67'	[2,3,5,6,7,8]	59,76%	94,90%	'newff_67'	[2,3,5,6,7,8]	40,48%	94,97%
'newff_68'	[2,3,5,6,7,9]	65,85%	94,90%	'newff_68'	[2,3,5,6,7,9]	61,90%	94,97%
'newff_69'	[2,3,5,6,8,9]	57,32%	94,80%	'newff_69'	[2,3,5,6,8,9]	52,38%	94,67%
'newff_70'	[2,3,5,7,8,9]	65,85%	94,77%	'newff_70'	[2,3,5,7,8,9]	69,05%	94,98%
'newff_71'	[2,3,6,7,8,9]	35,37%	95,00%	'newff_71'	[2,3,6,7,8,9]	19,05%	94,97%
'newff_72'	[2,4,5,6,7,8]	53,66%	94,92%	'newff_72'	[2,4,5,6,7,8]	42,86%	94,94%
'newff_73'	[2,4,5,6,7,9]	58,54%	94,89%	'newff_73'	[2,4,5,6,7,9]	71,43%	94,87%
'newff_74'	[2,4,5,6,8,9]	64,63%	94,90%	'newff_74'	[2,4,5,6,8,9]	61,90%	94,81%
'newff_75'	[2,4,5,7,8,9]	68,29%	94,93%	'newff_75'	[2,4,5,7,8,9]	69,05%	94,79%
'newff_76'	[2,4,6,7,8,9]	42,68%	94,91%	'newff_76'	[2,4,6,7,8,9]	23,81%	94,69%
'newff_77'	[2,5,6,7,8,9]	58,54%	94,92%	'newff_77'	[2,5,6,7,8,9]	52,38%	94,99%
'newff_78'	[3,4,5,6,7,8]	79,27%	94,67%	'newff_78'	[3,4,5,6,7,8]	78,57%	94,94%
'newff_79'	[3,4,5,6,7,9]	86,59%	94,54%	'newff_79'	[3,4,5,6,7,9]	90,48%	94,75%
'newff_80'	[3,4,5,6,8,9]	80,49%	94,67%	'newff_80'	[3,4,5,6,8,9]	85,71%	94,78%
'newff_81'	[3,4,5,7,8,9]	86,59%	94,89%	'newff_81'	[3,4,5,7,8,9]	90,48%	94,86%
'newff_82'	[3,4,6,7,8,9]	75,61%	94,79%	'newff_82'	[3,4,6,7,8,9]	73,81%	94,80%
'newff_83'	[3,5,6,7,8,9]	81,71%	94,87%	'newff_83'	[3,5,6,7,8,9]	80,95%	94,93%
'newff_84'	[4,5,6,7,8,9]	81,71%	94,76%	'newff_84'	[4,5,6,7,8,9]	80,95%	94,80%
Test 7 (7 features left out):							
Training Results:				Validation Results:			
Model ID	Features left out	T21	Normal	Model ID	Features left out	T21	Normal
'newff_1'	[1,2,3,4,5,6,7]	73,17%	94,98%	'newff_1'	[1,2,3,4,5,6,7]	61,90%	94,89%
'newff_2'	[1,2,3,4,5,6,8]	62,20%	94,75%	'newff_2'	[1,2,3,4,5,6,8]	61,90%	94,64%
'newff_3'	[1,2,3,4,5,6,9]	74,39%	94,61%	'newff_3'	[1,2,3,4,5,6,9]	78,57%	94,85%
'newff_4'	[1,2,3,4,5,7,8]	68,29%	94,93%	'newff_4'	[1,2,3,4,5,7,8]	57,14%	94,88%
'newff_5'	[1,2,3,4,5,7,9]	73,17%	94,84%	'newff_5'	[1,2,3,4,5,7,9]	78,57%	94,73%
'newff_6'	[1,2,3,4,5,8,9]	75,61%	94,70%	'newff_6'	[1,2,3,4,5,8,9]	76,19%	94,87%
'newff_7'	[1,2,3,4,6,7,8]	36,59%	94,79%	'newff_7'	[1,2,3,4,6,7,8]	28,57%	94,69%
'newff_8'	[1,2,3,4,6,7,9]	68,29%	94,88%	'newff_8'	[1,2,3,4,6,7,9]	64,29%	94,75%
'newff_9'	[1,2,3,4,6,8,9]	54,88%	94,94%	'newff_9'	[1,2,3,4,6,8,9]	54,76%	94,97%
'newff_10'	[1,2,3,4,7,8,9]	60,98%	94,60%	'newff_10'	[1,2,3,4,7,8,9]	69,05%	94,64%
'newff_11'	[1,2,3,5,6,7,8]	50,00%	94,94%	'newff_11'	[1,2,3,5,6,7,8]	33,33%	94,81%

'newff_12'	[1,2,3,5,6,7,9]	73,17%	94,77%		'newff_12'	[1,2,3,5,6,7,9]	71,43%	94,90%
'newff_13'	[1,2,3,5,6,8,9]	52,44%	94,90%		'newff_13'	[1,2,3,5,6,8,9]	64,29%	94,65%
'newff_14'	[1,2,3,5,7,8,9]	67,07%	94,94%		'newff_14'	[1,2,3,5,7,8,9]	69,05%	94,74%
'newff_15'	[1,2,3,6,7,8,9]	35,37%	94,47%		'newff_15'	[1,2,3,6,7,8,9]	23,81%	94,26%
'newff_16'	[1,2,4,5,6,7,8]	56,10%	94,88%		'newff_16'	[1,2,4,5,6,7,8]	33,33%	94,80%
'newff_17'	[1,2,4,5,6,7,9]	68,29%	94,94%		'newff_17'	[1,2,4,5,6,7,9]	64,29%	94,93%
'newff_18'	[1,2,4,5,6,8,9]	60,98%	94,99%		'newff_18'	[1,2,4,5,6,8,9]	50,00%	94,69%
'newff_19'	[1,2,4,5,7,8,9]	65,85%	94,95%		'newff_19'	[1,2,4,5,7,8,9]	71,43%	94,86%
'newff_20'	[1,2,4,6,7,8,9]	36,59%	94,95%		'newff_20'	[1,2,4,6,7,8,9]	19,05%	94,68%
'newff_21'	[1,2,5,6,7,8,9]	57,32%	94,90%		'newff_21'	[1,2,5,6,7,8,9]	47,62%	94,82%
'newff_22'	[1,3,4,5,6,7,8]	75,61%	95,00%		'newff_22'	[1,3,4,5,6,7,8]	76,19%	94,77%
'newff_23'	[1,3,4,5,6,7,9]	85,37%	94,83%		'newff_23'	[1,3,4,5,6,7,9]	92,86%	94,83%
'newff_24'	[1,3,4,5,6,8,9]	80,49%	93,59%		'newff_24'	[1,3,4,5,6,8,9]	88,10%	93,70%
'newff_25'	[1,3,4,5,7,8,9]	86,59%	94,76%		'newff_25'	[1,3,4,5,7,8,9]	90,48%	94,79%
'newff_26'	[1,3,4,6,7,8,9]	70,73%	94,78%		'newff_26'	[1,3,4,6,7,8,9]	80,95%	94,87%
'newff_27'	[1,3,5,6,7,8,9]	81,71%	94,78%		'newff_27'	[1,3,5,6,7,8,9]	76,19%	94,77%
'newff_28'	[1,4,5,6,7,8,9]	82,93%	94,95%		'newff_28'	[1,4,5,6,7,8,9]	73,81%	94,97%
'newff_29'	[2,3,4,5,6,7,8]	34,15%	94,99%		'newff_29'	[2,3,4,5,6,7,8]	23,81%	94,77%
'newff_30'	[2,3,4,5,6,7,9]	63,41%	94,91%		'newff_30'	[2,3,4,5,6,7,9]	59,52%	94,75%
'newff_31'	[2,3,4,5,6,8,9]	50,00%	94,89%		'newff_31'	[2,3,4,5,6,8,9]	52,38%	94,82%
'newff_32'	[2,3,4,5,7,8,9]	64,63%	94,54%		'newff_32'	[2,3,4,5,7,8,9]	52,38%	94,92%
'newff_33'	[2,3,4,6,7,8,9]	24,39%	94,55%		'newff_33'	[2,3,4,6,7,8,9]	0,00%	100,00%
'newff_34'	[2,3,5,6,7,8,9]	34,15%	94,84%		'newff_34'	[2,3,5,6,7,8,9]	35,71%	94,97%
'newff_35'	[2,4,5,6,7,8,9]	37,80%	94,95%		'newff_35'	[2,4,5,6,7,8,9]	19,05%	94,77%
'newff_36'	[3,4,5,6,7,8,9]	74,39%	94,79%		'newff_36'	[3,4,5,6,7,8,9]	83,33%	94,62%
Test 8 (8 features left out):								
Training Results:					Validation Results:			
Model ID	Features left out	T21	Normal		Model ID	Features left out	T21	Normal
'newff_1'	[1,2,3,4,5,6,7,8]	31,71%	94,98%		'newff_1'	[1,2,3,4,5,6,7,8]	21,43%	94,75%
'newff_2'	[1,2,3,4,5,6,7,9]	64,63%	94,91%		'newff_2'	[1,2,3,4,5,6,7,9]	64,29%	94,61%
'newff_3'	[1,2,3,4,5,6,8,9]	45,12%	98,33%		'newff_3'	[1,2,3,4,5,6,8,9]	52,38%	98,29%
'newff_4'	[1,2,3,4,5,7,8,9]	58,54%	96,26%		'newff_4'	[1,2,3,4,5,7,8,9]	66,67%	96,45%
'newff_5'	[1,2,3,4,6,7,8,9]	18,29%	94,88%		'newff_5'	[1,2,3,4,6,7,8,9]	9,52%	94,96%
'newff_6'	[1,2,3,5,6,7,8,9]	26,83%	94,87%		'newff_6'	[1,2,3,5,6,7,8,9]	21,43%	94,71%
'newff_7'	[1,2,4,5,6,7,8,9]	35,37%	94,42%		'newff_7'	[1,2,4,5,6,7,8,9]	19,05%	94,90%
'newff_8'	[1,3,4,5,6,7,8,9]	71,95%	94,90%		'newff_8'	[1,3,4,5,6,7,8,9]	80,95%	94,97%
'newff_9'	[2,3,4,5,6,7,8,9]	18,29%	94,88%		'newff_9'	[2,3,4,5,6,7,8,9]	11,90%	94,10%

Πείραμα 6

Στα επόμενα πειράματα χρησιμοποίησα τη βάση δεδομένων η οποία περιείχε τα πολλαπλάσια του μέσου όρου (Multiple of the median) στις τιμές των BhCG και PAPP-A. Ο ενδιάμεσος μέσος όρος χρησιμοποιείται για να δείξει την απόκλιση που έχει ένα αποτέλεσμα μιας εξέτασης από την διάμεση τιμή. Για παράδειγμα για τιμή 1 στο MoM σημαίνει ότι το επίπεδο δείκτη είναι ίσο με τη διάμεση τιμή. Σε αυτό το πείραμα είχαμε ένα κρυφό νευρώνα στα μοντέλα μας, σταθερό αριθμό εποχών στις 200 και εξετάσαμε τους νευρώνες τους οποίους μετακινήσαμε από 5 μέχρι 70, αυξάνοντας τους κατά 5 σε κάθε μοντέλο. Η συνάρτηση ενεργοποίησης του επιπέδου εισόδου είναι η tansig και η συνάρτηση ενεργοποίησης του κρυφού επιπέδου είναι η purelin.

Training Results				Validation Results			
Model ID	Neurons	T21	Normal	Model ID	Neurons	T21	Normal
'newff_1'	5	95,12%	92,93%	'newff_1'	5	97,62%	93,07%
'newff_2'	10	95,12%	94,51%	'newff_2'	10	92,86%	94,49%
'newff_3'	15	97,56%	92,55%	'newff_3'	15	83,33%	92,56%
'newff_4'	20	96,34%	94,80%	'newff_4'	20	83,33%	94,78%
'newff_5'	25	98,78%	94,82%	'newff_5'	25	76,19%	94,91%
'newff_6'	30	100,00%	94,72%	'newff_6'	30	88,10%	94,90%
'newff_7'	35	97,56%	94,73%	'newff_7'	35	64,29%	94,93%
'newff_8'	40	98,78%	94,53%	'newff_8'	40	54,76%	94,36%
'newff_9'	45	97,56%	94,99%	'newff_9'	45	71,43%	94,89%
'newff_10'	50	98,78%	95,00%	'newff_10'	50	66,67%	94,99%
'newff_11'	55	97,56%	94,73%	'newff_11'	55	83,33%	94,93%
'newff_12'	60	98,78%	94,90%	'newff_12'	60	66,67%	94,83%
'newff_13'	65	98,78%	94,98%	'newff_13'	65	71,43%	94,75%
'newff_14'	70	98,78%	94,80%	'newff_14'	70	69,05%	94,90%

Πείραμα 7

Και σε αυτό το πείραμα χρησιμοποίησα τα δεδομένα με τις MoM τιμές όπως στο προηγούμενο πείραμα κατασκευάζοντας τον ίδιο τύπο νευρωνικών δικτύων, αυτή τη φορά με 2 κρυφά επίπεδα εξετάζοντας για νευρώνες από 10 μέχρι 50 και στα δύο, αυξάνοντας τους κατά 5. Ο αριθμός των εποχών και των συναρτήσεων ενεργοποίησης παρέμειναν σταθερές σε 200, tansig και purelin αντίστοιχα.

Training Results				Validation Results			
Model ID	Neurons	T21	Normal	Model ID	Neurons	T21	Normal
newff_1	[10,10]	93,90%	94,68%	newff_1	[10,10]	90,48%	94,59%
newff_2	[15,15]	96,34%	94,96%	newff_2	[15,15]	85,71%	93,02%
newff_3	[20,20]	97,56%	94,80%	newff_3	[20,20]	85,71%	94,66%
newff_4	[25,25]	97,56%	94,88%	newff_4	[25,25]	73,81%	94,94%
newff_5	[30,30]	97,56%	94,86%	newff_5	[30,30]	90,48%	94,83%
newff_6	[35,35]	97,56%	94,71%	newff_6	[35,35]	64,29%	94,74%
newff_7	[40,40]	97,56%	94,66%	newff_7	[40,40]	69,05%	94,94%
newff_8	[45,45]	97,56%	94,70%	newff_8	[45,45]	88,10%	94,93%
newff_9	[50,50]	97,56%	94,96%	newff_9	[50,50]	95,24%	94,88%

Κεφάλαιο 6

Συμπεράσματα

6.1 Συμπεράσματα

6.1 Συμπεράσματα

Σε αυτό το κεφάλαιο θα αναφερθώ στα συμπεράσματα της μελέτης μου, στη σημασία κάποιων χαρακτηριστικών, θα σχολιάσω τα αποτελέσματα και κατά πόσον ήταν αναμενόμενα η όχι και θα συγκρίνω τα αποτελέσματα μου με αποτελέσματα άλλων εργασιών που έχω μελετήσει, πάνω στο ίδιο θέμα.

Στο σύνολο, παρουσίασα αποτελέσματα για 650 μοντέλα. Από όλα αυτά τα πειράματα υπήρχαν μοντέλα που έφεραν ικανοποιητικά αποτελέσματα με επιθυμητά ποσοστά επιτυχίας, αλλά και πολλά που αποτύγχαναν να δώσουν μια καλή κατηγοριοποίηση. Σαν πρώτο συμπέρασμα μπορώ να πω ότι τα πλείστα νευρωνικά μοντέλα είχαν ποσοστά ικανοποιητικά, γεγονός που κάνει τη χρήση Τεχνητών Νευρωνικών Δικτύων τουλάχιστο σαν πρώτη διάγνωση για χρωμοσωμικές ανωμαλίες κατά την εγκυμοσύνη, δυνατή. Υπάρχει φυσικά περιθώριο βελτίωσης των αποτελεσμάτων για ακόμα πιο ακριβείς απαντήσεις. Κάποια αποτελέσματα που είχα, επαλήθευσαν κάποια συμπεράσματα προηγούμενων μελετών που ασχολήθηκαν με το θέμα, ενώ κάποια άλλα δεν ήταν αναμενόμενα.

Στα μοντέλα με 2 κρυφά επίπεδα παρατηρήθηκαν καλύτερα αποτελέσματα σχετικά με τα μοντέλα ενός κρυφού επιπέδου. Πιο συγκεκριμένα είχαμε ποσοστά επιτυχίας γύρω στο 94% με 98% κατά την εκπαίδευση και γύρω στο 90% με 94% κατά την επαλήθευση. Στη συνέχεια, παρατηρήθηκε ότι σε πειράματα με μοντέλα 2 κρυφών επιπέδων ξεχώρισαν τα μοντέλα που είχαν 10 νευρώνες στα δύο κρυφά επίπεδα, έχοντας ποσοστά επιτυχίας κατά την εκπαίδευση γύρω στο 94% με 100% για φυσιολογικά περιστατικά και περιστατικά τρισωμίας 21 και στην επαλήθευση γύρω στο 94% με 97% για τις 2 κατηγορίες. Αυτά τα αποτελέσματα

είναι πολύ ενθαρρυντικά και μπορούν να ληφθούν σοβαρά υπόψη από τους γιατρούς, αν χρησιμοποιηθούν σε πραγματικά προβλήματα.

Στη συνέχεια δημιουργήθηκαν μοντέλα με συνδυασμούς χαρακτηριστικών εισόδου για να μελετηθούν τα αποτελέσματα που μπορούν να δώσουν κάποιοι συνδυασμοί οι οποίοι αγνοούν κάποια χαρακτηριστικά. Στα μοντέλα όπου το σύστημα αγνοούσε 1 χαρακτηριστικό φάνηκε ότι χωρίς την τριγλώχινια ροή, τα αποτελέσματα ήταν ψηλότερα από τα υπόλοιπα μοντέλα, πράγμα που σημαίνει ότι η τριγλώχινια ροή επηρεάζει το σύστημα αρνητικά και δεν έχει καθοριστική σημασία.

Στα μοντέλα που χρησιμοποιούσαν 1 χαρακτηριστικό εισόδου, ξεχώρισαν τα αποτελέσματα στα οποία χρησιμοποιήθηκε μόνο η αυχενική διαφάνεια με σχετικά καλά ποσοστά πρόβλεψης. Αυτό το γεγονός κάνει την αυχενική διαφάνεια καθοριστικής σημασίας για την πρόβλεψη χρωμοσωμικής ανωμαλίας. Φάνηκε επίσης ότι καθοριστικής σημασίας είναι και η ροή φλεβώδους πόρου, η οποία όταν χρησιμοποιήθηκε μόνη της σαν χαρακτηριστικό εισόδου έφερε επίσης καλά αποτελέσματα πρόβλεψης. Αυτά τα 2 χαρακτηριστικά επιβεβαίωσαν την καθοριστική σημασία που έχουν, εφόσον είχαν ψηλά αποτελέσματα και όταν χρησιμοποιήθηκαν ως συνδυασμός χαρακτηριστικών εισόδου. Παρατηρήθηκαν ακόμα υψηλότερα αποτελέσματα όταν προσομοιώθηκε μοντέλο που περιείχε σαν είσοδο αυτά τα 2 χαρακτηριστικά σε συνδυασμό με την χοριακή γονιαδοτροπίνη, δείχνοντας έτσι, εκτός από την σημαντικότητα της αυχενικής διαφάνειας και της ροής φλεβώδους πόρου, τη μεγάλη σημασία της χοριακής γονιαδοτροπίνης. Τα πειράματα αυτά επιβεβαιώνουν επίσης κάποια από τα συμπεράσματα της εργασίας [14] τα οποία έλεγαν ότι η αυχενική διαφάνεια είναι ο πιο σημαντικός παράγοντας πρόβλεψης, ενώ το κεφαλουραίο μήκος ο λιγότερο σημαντικός, εφόσον στα δικά μου πειράματα όταν χρησιμοποιήθηκε μόνο του στην είσοδο δεν έφερε ικανοποιητικά αποτελέσματα ενώ στα πειράματα που αγνοήθηκε δεν επηρέασε ιδιαίτερα τα αποτελέσματα.

Αυτή η εργασία [8] συμπέρανε ότι η ηλικία της μητέρας σε συνδυασμό με την αυχενική διαφάνεια και τη ροή φλεβώδους πόρου είχε καθοριστικά αποτελέσματα σε μια πρώτης φάσης εξέταση με σκοπό την σωστή πρόβλεψη τυχών τρισωμίας 21 στο έμβρυο. Στα πειράματα μου, επιβεβαιώνεται η πρόταση αυτή έχοντας ποσοστά επιτυχίας γύρω στο 85% με 95% στην φάση εκπαίδευσης και 80% με 95% στην φάση επαλήθευσης, χρησιμοποιώντας αυτόν τον συνδυασμό δεδομένων. Τέλος, αυτή η εργασία [18] παρουσιάζει αποτελέσματα με ποσοστό επιτυχίας 87% όταν χρησιμοποιηθεί συνδυασμός της αυχενικής διαφάνειας, της χοριακής γονιαδοτροπίνης και της πρωτεΐνης A στην είσοδο. Και εδώ τα πειράματα μου επαληθεύουν την επιτυχία του συνδυασμού αυτού φέρνοντας εξίσου ψηλά αποτελέσματα.

Αξιοσημείωτη παρατήρηση είναι ότι ο συνδυασμός της αυχενικής διαφάνειας με την πρωτεΐνη Α, την ροή φλεβώδους πόρου, τον καρδιακό σφυγμό και την ύπαρξη ή όχι ρινικού οστού είχε πολύ ψηλά ποσοστά πρόβλεψης που ξεπερνούν το 93% σε εκπαίδευση και επαλήθευση. Αυτό το αποτέλεσμα μας δείχνει ότι ο συνδυασμός των πιο πάνω χαρακτηριστικών είναι αρκετός για να πάρουμε μια πρώτη ιδέα για την υγεία του εμβρύου.

Τα επόμενα πειράματα που έκανα χρησιμοποίησαν τη βάση δεδομένων με τιμές MoM. Τα καλύτερα αποτελέσματα παρατηρήθηκαν στα μοντέλα με 5 και 10 νευρώνες στο κρυφό επίπεδο με ποσοστά επιτυχίας 92% με 94% για τα φυσιολογικά περιστατικά και 93% με 97% για την τρισωμία 21. Τα αποτελέσματα στην φάση εκπαίδευσης δεν διαφέρουν πολύ από τα αποτελέσματα που είχα στα ίδια μοντέλα όπου χρησιμοποίησα στην είσοδο τα κανονικά δεδομένα. Στην φάση επαλήθευσης, όμως, παρατηρήθηκαν βελτιωμένα αποτελέσματα με τις τιμές MoM, γεγονός που δεν ήταν αναμενόμενο, εφόσον σε προηγούμενες εξετάσεις χρησιμοποιώντας τις τιμές MoM αποδείχτηκε ότι έχουν χειρότερα αποτελέσματα προβλέψεων.

Περίληψη Συμπερασμάτων

Η αυχενική διαφάνεια μαζί με την ροή φλεβώδους πόρου σαν χαρακτηριστικά εισόδου έφεραν τα πιο κάτω ποσοστά επιτυχίας σε εκπαίδευση και επαλήθευση, τα οποία είναι πολύ ικανοποιητικά αν σκεφτούμε ότι χρησιμοποιούμε μόνο 2 χαρακτηριστικά εισόδου. Η καθοριστική σημασία που έχουν αυτά τα δύο χαρακτηριστικά είναι εμφανής:

Αυχενική διαφάνεια + Ροή φλεβώδους πόρου	T21	Normal
Training Results	85,37%	94,83%
Validation Results	92,86%	94,83%

Στον επόμενο πίνακα φαίνεται ότι η χοριακή γονιαδοτροπίνη είναι επίσης μεγάλης σημασίας, εφόσον όταν χρησιμοποιήθηκε μαζί με τα προηγούμενα 2 χαρακτηριστικά, το δίκτυο έφερε ακόμη καλύτερα αποτελέσματα:

Αυχενική διαφάνεια + Ροή φλεβώδους πόρου + Χοριακή γονιαδοτροπίνη	T21	Normal
Training Results	91,46%	95,00%
Validation Results	90,48%	94,93%

Στη συνέχεια παρουσιάζονται σε μορφή πίνακα τα αποτελέσματα από τα πειράματα για τη σημασία των χαρακτηριστικών και γίνεται μια σύγκριση με τα αποτελέσματα που είχε το νευρωνικό δίκτυο στο οποίο δώσαμε σαν είσοδο όλα τα 9 χαρακτηριστικά που εξετάζονται. Το δίκτυο που κρατώ σαν σημείο αναφοράς για τις συγκρίσεις είναι αυτό που είχε ακριβώς την ίδια αρχιτεκτονική και παραμέτρους που χρησιμοποίησα στα πειράματα συνδυασμών χαρακτηριστικών εισόδου, και η σύγκριση γίνεται με βάση τα ποσοστά επιτυχίας κατά τη φάση επαλήθευσης. Το εν λόγω δίκτυο είχε ποσοστά επιτυχίας 83,33% για πρόβλεψη Τρισωμίας 21 και 94,67% για πρόβλεψη φυσιολογικών περιστατικών. Στην τελευταία στήλη των πινάκων αυτών φαίνεται η ποσοστιαία αύξηση ή μείωση επιτυχίας, όταν είχαμε σαν είσοδο τον συνδυασμό χαρακτηριστικών που φαίνεται αριστερότερα.

Πίνακας επεξήγησης χαρακτηριστικών

CRL	Κεφαλουραίο μήκος
NT	Αυχενική διαφάνεια
PAPP-A	Πρωτεΐνη Α
b-hCG	Χοριακή γοναδιοτροπίνη
H. Rate	Καρδιακός Σφυγμός
NB	Ρινικό οστό
TF	Τριγλώχινα ροή
DV PI	Ροή φλεβώδους πόρου
MA	Ηλικία μητέρας

CRL	NT	PAPP-A	b-hCG	H. Rate	NB	TF	DV PI	MA	Validation Results		Result Comparisons	
									T21	Normal		
	✓	✓	✓	✓	✓	✓	✓	✓	92,86%	94,98%	↑	10%
✓	✓		✓	✓	✓	✓	✓	✓	95,24%	94,59%	↑	12%
✓	✓	✓	✓	✓	✓		✓	✓	97,62%	94,73%	↑	14%
	✓		✓		✓	✓	✓	✓	97,62%	94,45%	↑	14%
	✓	✓		✓	✓		✓		97,62%	94,77%	↑	14%
	✓		✓		✓		✓		95,24%	94,56%	↑	12%
	✓		✓				✓		90,48%	94,93%	↑	7%

Στον πιο πάνω πίνακα παρουσιάζονται οι συνδυασμοί χαρακτηριστικών οι οποίοι ξεχώρισαν για τα υψηλά ποσοστά επιτυχίας που έφεραν. Μπορεί κανείς να προσέξει ποιοι συνδυασμοί μπορούν να επιτύχουν σε διαγνώσεις τις Τρισωμίας 21.

Η χαμηλή σημασία του Κεφαλουραίου μήκους, της Τριγλώχινας ροής και της Πρωτεΐνης Α, αν αγνοηθούν ξεχωριστά, μπορεί να συμπεραθεί, εφόσον στα πειράματα που έγιναν αυτές οι δοκιμές είχαμε καλύτερα αποτελέσματα. Μπορεί επίσης να θεωρηθεί ότι κάθε ένα από αυτά τα 3 χαρακτηριστικά προκαλεί μια «σύγχυση» στο δίκτυο. Η παράληψη του Κεφαλουραίου μήκους αύξησε τα ποσοστά επιτυχίας της πρόβλεψης της Τρισωμίας 21 στην επαλήθευση κατά 10%, η παράληψη της Πρωτεΐνης Α αύξησε τα ποσοστά κατά 12% και η παράληψη της Τριγλώχινας ροής κατά 14%.

Η επόμενη σημαντική παρατήρηση είναι ότι ο συνδυασμός της Αυχενικής διαφάνειας, με την Πρωτεΐνη Α, τη Ροή φλεβώδους πόρου, τον Καρδιακό σφυγμό του εμβρύου και το Ρινικό οστό στην είσοδο έφερε εντυπωσιακά αποτελέσματα, ιδιαίτερα στην φάση επαλήθευσης, γεγονός που κάνει τον συνδυασμό αυτό καθοριστικό σε πρώιμες διαγνώσεις για Τρισωμία 21.

Περαιτέρω επιβεβαίωση του συμπεράσματος που σχολίασα πριν, ότι δηλαδή η Αυχενική διαφάνεια και η Ροή φλεβώδους πόρου είναι τεράστιας σημασίας, έρχεται αν παρατηρήσουμε στον πίνακα, ότι σε όλα τα μοντέλα που είχαμε αύξηση της επιτυχίας, αυτά τα 2 χαρακτηριστικά ήταν μέρος της εισόδου.

CRL	NT	PAPP-A	b-hCG	H. Rate	NB	TF	DV PI	MA	Validation Results		Result Comparisons	
									T21	Normal		
✓			✓	✓	✓			✓	47,62%	94,87%	↓	-36%
✓		✓		✓				✓	38,10%	94,83%	↓	-45%
✓				✓				✓	23,81%	94,83%	↓	-60%
✓			✓	✓					19,05%	94,97%	↓	-64%
✓		✓		✓					23,81%	94,69%	↓	-60%
				✓				✓	28,57%	94,69%	↓	-55%
			✓	✓					23,81%	94,26%	↓	-60%
		✓		✓					19,05%	94,68%	↓	-64%
✓								✓	23,81%	94,77%	↓	-60%

Σε αυτόν τον πίνακα παρουσιάζονται τα μοντέλα που είχαν αρνητικά αποτελέσματα και μείωσαν κατακόρυφα την επιτυχία προβλέψεων. Φαίνονται ξεκάθαρα ποιοι συνδυασμοί χαρακτηριστικών δεν μπορούν σε καμία περίπτωση να δώσουν μια καλή απάντηση στο πρόβλημα. Αξιοσημείωτο πάντως, είναι το γεγονός ότι η ολική παράληψη του συνδυασμού της Αυχενικής διαφάνειας της Τριγλώχινας ροής και της Ροής φλεβώδους πόρου φαίνεται να είναι η κυριότερη αιτία της αποτυχίας αυτών των μοντέλων.

Στα 2 τελευταία πειράματα με τις τιμές MoM, παρουσιάστηκαν κατά τη φάση της επαλήθευσης καλύτερα αποτελέσματα σε σχέση με τα Raw data κατά ένα ποσοστό της τάξεως του 5%, γεγονός που δεν ήταν αναμενόμενο. Κατά τη φάση εκπαίδευσης δε φάνηκε ιδιαίτερη διαφορά στα ποσοστά επιτυχίας μεταξύ των δύο βάσεων δεδομένων. Φαίνεται η σύγκριση στον πιο κάτω πίνακα, λαμβάνοντας υπόψη ένα μέσο όρο των ποσοστών επιτυχίας όλων των αποτελεσμάτων για κάθε ένα από τα 2 πειράματα.

Πείραμα MoM

	Training Results			Validation Results	
Average Results	T21	Normal		T21	Normal
Πείραμα 1	98,03%	94,62%		74,73%	94,63%
Πείραμα 2	97,02%	94,80%		82,54%	94,62%

Πείραμα Raw Data

	Training Results			Validation Results	
Average Results	T21	Normal		T21	Normal
Πείραμα 1	99,06%	94,73%		69,41%	94,74%
Πείραμα 2	98,10%	94,82%		76,46%	94,82%

Κεφάλαιο 7

Θέματα Μελλοντικής Μελέτης

7.1 Θέματα μελλοντικής μελέτης

7.1 Θέματα μελλοντικής μελέτης

Σαν μελλοντική εργασία πάνω στο θέμα που ασχολήθηκα, σημαντική βελτίωση μπορεί να γίνει στον αλγόριθμο που φτιάχτηκε για να δημιουργεί διάφορα μοντέλα νευρωνικών δικτύων με διαφορετικές αρχιτεκτονικές κάθε φορά. Ο αλγόριθμος, μπορεί να εξελιχτεί έτσι ώστε να υποστηρίζει και άλλους αλγόριθμους μηχανικής μάθησης, όπως Support Vector Machines και K-nearest Neighbor.

Επίσης, τα μοντέλα που φτιάχτηκαν σε αυτή την μελέτη είναι μεν πολλά σε αριθμό, αλλά μπορούν συμπληρωματικά να υλοποιηθούν τεχνικές ομαδοποίησης δεδομένων όπως Self-organizing Maps και K-means για αρχικοποίηση των βάσεων εκπαίδευσης και επαλήθευσης.

Στην εργασία αυτή, δεν μελετήθηκαν οι υπόλοιπες τρισωμίες όπως T13, T18, Triploidy και Turner. Ένα θέμα μελλοντικής εργασίας θα μπορούσε να ήταν και η μελέτη και υπόλοιπων χρωμοσωμικών ανωμαλιών.

Αναφορές

- [1] M. Alvarez, “Computational Intelligence Techniques for Predicting Earthquakes”, Proceeding, HAIS'11 Proceedings of the 6th international conference on Hybrid artificial intelligent systems, Vol. part 2, pp 287-294, 2011.
- [2] B. Andrew, A. B. Sparks, T. Eric, et al. “Selective analysis of cell-free DNA in maternal blood for evaluation of fetal trisomy”, Prenatal Diagnosis Vol. 32, Issue 1, pp 3-9, 2012
- [3] M. Divsalar, A. K. Firouzabadi, M. Sadeghi, et al. “Artificial Intelligence towards the prediction of business failure via computational intelligence techniques”, Expert Systems Vol. 28, Issue 3, pp 209-226, 2011.
- [4] M. Ehrich, C. Deciu, T. Zwiefelhofer, et al. “Noninvasive detection of fetal trisomy 21 by sequencing of DNA in maternal blood: a study in a clinical setting”, American Journal of Obstetrics and Gynecology, Vol. 204, Issue 3, pp 205.e1–205.e11, 2011.
- [5] Fetal Medicine Foundation, Available at <http://www.fetalmedicine.com/fmf/> (ανακτήθηκε στις 10 Ιανουαρίου 2013).
- [6] Harvard Medical School (Amniocentesis), Available at <http://www.health.harvard.edu/diagnostic-tests/amniosentesis.htm> (ανακτήθηκε στις 25 Φεβρουαρίου 2013).
- [7] J. H. Holland, “Computer programs that evolve in ways that resemble natural selection can solve complex problems even their creators do not fully understand”, Scientific American, Vol. 267, pp 66-72, 1992.
- [8] K. O. Kagan, I. Staboulidou, J. Cruz, et al. “Two-stage first-trimester screening for trisomy 21 by ultrasound assessment and biochemical testing”, Ultrasound in Obstetrics and Gynecology, Vol. 36, Issue 5, pp 542-547, 2010.
- [9] O. Kulhanek, Seminar on b-value, Dept. of Geophysics, Charles University, Prague, December 2005.
- [10] K. Lee and W. S. Yang., “Historical seismicity of Korea, Bulletin of the Seismological Society of America”, Vol. 96, pp 846-855, 2006.

- [11] M. C. Macintosh, B. Brambati, T. Chard, et al. "Predicting fetal chromosome anomalies in the first trimester using pregnancy associated plasma protein-A: a comparison of statistical methods", *Methods of Information in Medicine*, Vol. 32, pp 175-179, 1993.
- [12] N. Maiz, C. Valencia, K.O. Kagan, et al. "Ductus venosus Doppler in screening for trisomies 21, 18 and 13 and Turner syndrome at 11-13 weeks of gestation", *Ultrasound in Obstetrics and Gynecology*, Vol. 33, Issue 5, pp 512-517, 2009.
- [13] Maternal serum screening, Available at http://www.betterhealth.vic.gov.au/bhcv2/bhcarticles.nsf/pages/Maternal_serum_screening (ανακτήθηκε στις 20 Ιανουαρίου 2013).
- [14] A. C. Neocleous, K. H. Nicolaides, A. Syngelaki, et al. "Artificial Neural Networks to Investigate the Importance and the Sensitivity to Various Parameters used for the Prediction of Chromosomal Abnormalities", *Artificial Intelligence Applications and Innovations*, Vol. 382, pp 46-55, 2012.
- [15] C. K. Neocleous, K. H. Nikolaides, K. C. Neokleous, et al. "Artificial Neural Networks to Investigate the Significance of PAPP-A and b-hCG for the Prediction of Chromosomal Abnormalities", *International Joint Conference on Neural Networks, IJCNN*, pp 1955 - 1958 , 2011.
- [16] C. N. Neocleous, K. Nikolaides, K. Neokleous, et al. "First Trimester Diagnosis of Trisomy-21 Using Artificial Neural Networks", *XII Mediterranean Conference on Medical and Biological Engineering and Computing*, Vol. 29, pp 580-583, 2010.
- [17] New Down Syndrome Test Is Accurate, Available at <http://specialpost.com/tag/digital-analysis-of-selected-regions/> (ανακτήθηκε στις 25 Φεβρουαρίου 2013).
- [18] K. H. Nicolaides, "Nuchal translucency and other first-trimester sonographic markers of chromosomal abnormalities", *American Journal of Obstetrics and Gynecology*, Vol. 191, Issue 1, pp 45-67, 2004.

[19] D. Pinkel, J. Landegent, C. Collins, et al. "Fluorescence in situ hybridization with human chromosome-specific libraries", Proceedings of the National Academy of Sciences USA, Vol. 85, pp 9138-9142, 1988.

[20] US National Library of Medicine (Chorionic villus sampling), Available at <http://www.nlm.nih.gov/medlineplus/ency/article/003406.htm> (ανακτήθηκε στις 25 Φεβρουαρίου 2013).

[21] L.A. Zadeh, Biography, Available at <http://www.eecs.berkeley.edu/Faculty/Homepages/zadeh.html> (ανακτήθηκε στις 10 Ιανουαρίου 2013).

[22] X. Zhang, The delta rule and learning for min-max neural networks, IEEE World Congress on Computational Intelligence, Vol. 1, pp 38-43, 1994.

[23] Ι. Λοΐζου, Μελέτη σημαντικότητας των χαρακτηριστικών εισόδου σε συστήματα υπολογιστικής νοημοσύνης για διάγνωση στην Ιατρική, Ατομική Διπλωματική Εργασία, Τμήμα Πληροφορικής, Πανεπιστήμιο Κύπρου, Μάιος 2012.

[24] Φ. Νεοφύτου, Συστήματα υπολογιστικής νοημοσύνης για διάγνωση στην Ιατρική: Η περίπτωση του συνδρόμου Down, Ατομική Διπλωματική Εργασία, Τμήμα Πληροφορικής, Πανεπιστήμιο Κύπρου, Μάιος 2011.

Παράρτημα Α

Main Code:

```
clear;clc
%% SELECT PROJECT FOLDER AND MAKE DIRECTORIES
input('Please select the folder of the project (Press enter): ')
cd(uigetdir)
clc
mkdir('Models')
%% GET DATABASE
[id,data,data2,targets,classes_numbers]=Get_database;
%% INPUT PARAMETERS
in_pars=input_parameters;
%% CROSS VALIDATION
Xval=crossvalidation(data2,targets,id);
%% FEATURE TEST
[All_results_f_tst_1cmb,All_results_f_tst,reply_feature_tst]=feature_test(X
val,in_pars,classes_numbers);
clearvars -except id data data2 targets classes_numbers in_pars Xval
All_results_f_tst_1cmb All_results_f_tst reply_feature_tst
%% NORMAL RUN
All_results=normal_run(Xval,in_pars,classes_numbers,reply_feature_tst);
clearvars -except id data data2 targets classes_numbers in_pars Xval
All_results Xval All_results_f_tst_1cmb All_results_f_tst
```

Create Neural Network:

```
function
[T_results,T_out,V_results,V_out]=create_NN_no_lr_mom(T,T_targ,T_class_popu
lation,V,V_targ,V_class_population,Neurons,Transfer_fn,Epochs,show_gui,clas
ses_numbers)
In_layer_TFns=Transfer_fn{1};
Layers_TFns=Transfer_fn{2};

T_results{1,1}='Network';
T_results{2,1}='Feed forward Back propagation';
T_results{4,1}='Model ID';
T_results{4,2}='Epochs';
T_results{4,3}='Input Transfer fn';

for i_lrs_tf=1:length(Layers_TFns)
    T_results{4,3+i_lrs_tf}=[ 'Layer_' int2str(i_lrs_tf) 'T.Fn' ];;
end
T_results{4,4+i_lrs_tf}='Neurons';
T_results{4,5+i_lrs_tf}='Threshold';
for icl=1:length(classes_numbers)
    T_results{4,5+i_lrs_tf+icl}=cellstr(['class' int2str(icl) ]);
end

V_results{1,1}='Network';
V_results{2,1}='Feed forward Back propagation';
V_results{4,1}='Model ID';
V_results{4,2}='Epochs';
V_results{4,3}='Input Transfer fn';
for i_lrs_tf=1:length(Layers_TFns)
    V_results{4,3+i_lrs_tf}=[ 'Layer_' int2str(i_lrs_tf) 'T.Fn' ];;
end
```

```

V_results{4,4+i_lrs_tf}='Neurons';
V_results{4,5+i_lrs_tf}='Threshold';
for icl=1:length(classes_numbers)
    V_results{4,5+i_lrs_tf+icl}=cellstr(['class' int2str(icl) ]);
end

iter=0;

% Neurons in hidden layer loops.
for ineurons=1:length(Neurons{1})
    % Input layer's transfer function loop
    for iIn_trans_fn=1:length(In_layer_TFns)
        % Layers transfer functions loop
        for ilayers_fns=1:length(Transfer_fn{2})
            % Epochs loop
            for iepochs=1:length(Epochs)
                clc
                iter=iter+1;
                disp(['Current iteration: ' int2str(iter)])
                %%% CREATE NEURAL NETWORK
                for ilayers=1:length(Layers_TFns)

neurons2net(ilayers)=Neurons{ilayers}(ineurons);
                end
                net = newff(T,T_targ,neurons2net);
                net.layers{1}.transferFcn =
In_layer_TFns{iIn_trans_fn};
                for ilayers2=1:length(Layers_TFns)
                    net.layers{ilayers2+1}.transferFcn =
Layers_TFns{ilayers2};

                end
                net.trainParam.epochs=Epochs(iepochs);
                net.trainParam.goal=1e-100;
                net.divideFcn='';
                net.trainParam.showWindow=show_gui;
                rand('seed',2);
                net=init(net);
                net=train(net,T,T_targ);
                model_id=['newff_' int2str(iter)];
                save (model_id,'net')

                % VALIDATION PROCESS
                T_out{1,1}='Training outputs';
                T_out{2,iter}=model_id;
                T_out{3,iter}=sim(net,T)';

                V_out{1,1}='Validation outputs';
                V_out{2,iter}=model_id;
                V_out{3,iter}=sim(net,V)';

[T_results1{1,iter},T_thr]=get_results(T_out{3,iter},T_targ,T_class_populat
ion,classes_numbers);

[V_results1{1,iter},V_thr]=get_results(V_out{3,iter},V_targ,V_class_populat
ion,classes_numbers);

                % CREATE TABLES

                % TRAINING
                % 1. Model ID

                T_results{iter+4,1}=model_id;

```

```

        % 2. Epochs
        T_results{iter+4,2}=num2str(Epochs(iePOCHs));
        % 5. Input layer transfer function

T_results{iter+4,3}=In_layer_TFns{iIn_trans_fn};
        % 6+n. Hidden layer(s) transfer function(s)
        for i_lrs_tf=1:length(Layers_TFns)

T_results{iter+4,3+i_lrs_tf}=Layers_TFns{ilayers2};
        end
        % 6+n+1. Neurons
        T_results{iter+4,4+i_lrs_tf}=neurons2net;
        % Print threshold
        T_results{iter+4,5+i_lrs_tf}=T_thr;
        % Print results for each class
        for i=1:length(classes_numbers)

T_results{iter+4,5+i_lrs_tf+i}=T_results1{1,iter}(i);
        end

        % VALIDATION
        % 1. Model ID
        V_results{iter+4,1}=model_id;
        % 2. Epochs
        V_results{iter+4,2}=num2str(Epochs(iePOCHs));
        % 5. Input layer transfer function

V_results{iter+4,3}=In_layer_TFns{iIn_trans_fn};
        % 6+n. Hidden layer(s) transfer function(s)
        for i_lrs_tf=1:length(Layers_TFns)

V_results{iter+4,3+i_lrs_tf}=Layers_TFns{ilayers2};
        end
        % 6+n+1. Neurons
        V_results{iter+4,4+i_lrs_tf}=neurons2net;
        % Print threshold
        V_results{iter+4,5+i_lrs_tf}=V_thr;

        % Print results for each class
        for i=1:length(classes_numbers)

V_results{iter+4,5+i_lrs_tf+i}=V_results1{1,iter}(i);
        end

        end

    end

end
end
end
end

```

Create Neural Network with Feature Test:

```
function
[T_results,T_out,V_results,V_out]=create_NN_no_lr_mom_ftst(T,T_targ,T_class
_population,V,V_targ,V_class_population,Neurons,Transfer_fn,Epochs,show_gui
,classes_numbers,combos)
In_layer_TFns=Transfer_fn{1};
Layers_TFns=Transfer_fn{2};

T_results{1,1}='Network';
T_results{2,1}='Feed forward Back propagation';
T_results{4,1}='Model ID';
T_results{4,2}='Features left out';
T_results{4,3}='Epochs';
T_results{4,4}='Input Transfer fn';

for i_lrs_tf=1:length(Layers_TFns)
    T_results{4,4+i_lrs_tf}={['Layer_' int2str(i_lrs_tf) 'T.Fn']};
end
T_results{4,5+i_lrs_tf}='Neurons';
T_results{4,6+i_lrs_tf}='Threshold';
for icl=1:length(classes_numbers)
    T_results{4,6+i_lrs_tf+icl}=cellstr(['class' int2str(icl) ]);
end

V_results{1,1}='Network';
V_results{2,1}='Feed forward Back propagation';
V_results{4,1}='Model ID';
V_results{4,2}='Features left out';
V_results{4,3}='Epochs';
V_results{4,4}='Input Transfer fn';
for i_lrs_tf=1:length(Layers_TFns)
    V_results{4,4+i_lrs_tf}={['Layer_' int2str(i_lrs_tf) 'T.Fn']};
end
V_results{4,5+i_lrs_tf}='Neurons';
V_results{4,6+i_lrs_tf}='Threshold';
for icl=1:length(classes_numbers)
    V_results{4,6+i_lrs_tf+icl}=cellstr(['class' int2str(icl) ]);
end

iter=0;

% Neurons in hidden layer loops.
for ineurons=1:length(Neurons{1})
    % Input layer's transfer function loop
    for iIn_trans_fn=1:length(In_layer_TFns)
        % Layers transfer functions loop
        for ilayers_fns=1:length(Transfer_fn{2})
            % Epochs loop
            for iepochs=1:length(Epochs)

                %%% CREATE NEURAL NETWORK
                for ilayers=1:length(Layers_TFns)

neurons2net(ilayers)=Neurons{ilayers}(ineurons);
                end

                for icomps2=1:length(combos(:,1))
                    clc
```

```

        iter=iter+1;
        disp(['Current iteration: ' int2str(iter)])
        T2=T;
        V2=V;
        % Create feature test

T2=T2(setdiff(1:size(T2,1),combos(icomps2,:)),:);

V2=V2(setdiff(1:size(V2,1),combos(icomps2,:)),:);
        net = newff(T2,T_targ,neurons2net);
        net.layers{1}.transferFcn =

In_layer_TFns{iIn_trans_fn};

        for ilayers2=1:length(Layers_TFns)
            net.layers{ilayers2+1}.transferFcn =

Layers_TFns{ilayers2};

        end
        net.trainParam.epochs=Epochs(iePOCHS);
        net.trainParam.goal=1e-100;
        net.divideFcn='';
        net.trainParam.showWindow=show_gui;
        rand('seed',2);
        net = init(net);
        net = train(net,T2,T_targ);
        model_id=['newff_' int2str(iter)];
        save (model_id,'net')
        % VALIDATION PROCESS

        T_out{1,1}='Training outputs';
        T_out{2,iter}=model_id;
        T_out{3,iter}=sim(net,T2)';

        V_out{1,1}='Validation outputs';
        V_out{2,iter}=model_id;
        V_out{3,iter}=sim(net,V2)';

[T_results1{1,iter},T_thr]=get_results(T_out{3,iter},T_targ,T_class_population,classes_numbers);

[V_results1{1,iter},V_thr]=get_results(V_out{3,iter},V_targ,V_class_population,classes_numbers);

        % CREATE TABLES

        % TRAINING
        % 1. Model ID
        T_results{iter+4,1}=model_id;
        % 2. Features to remove
        T_results{iter+4,2}=combos(icomps2,:);
        % 2. Epochs

T_results{iter+4,3}=num2str(EPOCHS(iePOCHS));
        % 5. Input layer transfer function

T_results{iter+4,4}=In_layer_TFns{iIn_trans_fn};
        % 6+n. Hidden layer(s) transfer function(s)
        for i_lrs_tf=1:length(Layers_TFns)

T_results{iter+4,4+i_lrs_tf}=Layers_TFns{ilayers2};
        end
        % 6+n+1. Neurons
        T_results{iter+4,5+i_lrs_tf}=neurons2net;
        % Print threshold
        T_results{iter+4,6+i_lrs_tf}=T_thr;
        % Print results for each class
        for i=1:length(classes_numbers)

```


Cross Validation:

```
function Xval=crossvalidation(data,targets,id)
reply1=input('Do you want to create cross validation sets? (y/n): ','s');
clc
if reply1=='y'
    reply2=input('How many folds do you want to create?: ');
    initial_folds=double(reply2);

    clc
    reply_folds_mdls=input('Create models for the following folds: (number
or "All"): ','s');
    if strcmp(reply_folds_mdls,'All')
        cross_validation=1:double(reply2);
    else
        cross_validation=str2num(reply_folds_mdls);
    end

    cd Models
    for ifold_folders=1:length(cross_validation)
        folder_fold=['Fold' int2str(cross_validation(ifold_folders))];
        mkdir(folder_fold)
    end
    cd ..

    unique1=unique(targets);
    number_of_classes=length(unique1);
    class_bin=zeros(number_of_classes,length(targets));

    % Create binary output for each class. Will be later used for
separating
    % the several classes.
    for iclass=1:length(targets)
        class_bin(:,iclass)=targets(iclass)==unique1;
    end

    % Separate n classes into a string vector. Each cell contains the data
% of the n'th class.
    class=cell(1,number_of_classes);
    for iclass_vec=1:number_of_classes
        class{iclass_vec}=data(:,find(class_bin(iclass_vec,:)));
        class_ids{iclass_vec}=id(find(class_bin(iclass_vec,:)));
    end

    % For each class, create a vector with "randperm" (see help).
    choose_sets_T_and_V=cell(1,number_of_classes);
    for ichoose_sets=1:number_of_classes

choose_sets_T_and_V{ichoose_sets}=randperm(length(class{1,ichoose_sets}));

        for ifoldsXV=1:length(cross_validation)

h=circshift(choose_sets_T_and_V{ichoose_sets}',floor(length(targets)/length
(cross_validation)));

choose_sets_T_and_V1{cross_validation(ifoldsXV),ichoose_sets}=h';
            end
        end
    end
```

```

    % Prepare index vector for next validation

    for ifolds=1:length(cross_validation)
        % sets_indices{cross_validation(ifolds)}=choose_sets_T_and_V;

        for i=1:number_of_classes
            l=length(class{1,i});

            trainInd{i}=choose_sets_T_and_V1{cross_validation(ifolds),i}(1:floor(1*(100-
            1/initial_folds*100)/100));

            valInd{i}=choose_sets_T_and_V1{cross_validation(ifolds),i}(floor(1*(100-
            1/initial_folds*100)/100)+1:l);

            [train_id{i},val_id{i}] =
            divideind(class_ids{1,i}',trainInd{i},valInd{i});
            [trainP{i},valP{i}] =
            divideind(class{1,i},trainInd{i},valInd{i});
            training_class_populations(i)=length(trainP{i}(1,:));
            evaluation_class_populations(i)=length(valP{i}(1,:));

            targ_tr(1:length(trainP{1,i}(1,:)))=unique1(i);
            training_targets1{i}=targ_tr;
            targ_ev(1:length(valP{1,i}(1,:)))=unique1(i);
            evaluation_targets1{i}=targ_ev;
            clear targ_tr;
            clear targ_ev;
        end
        % Training OUTS
        training{cross_validation(ifolds)}=cell2mat(trainP);

        training_targets{cross_validation(ifolds)}=cell2mat(training_targets1);
        training_ids{cross_validation(ifolds)}=cell2mat(train_id);
        training_indices=cell2mat(trainInd);
        training_indices=training_indices';
        indices{1,1}='Training_indices';
        indices{2,1}=training_indices;
        % Evaluation OUTS

        evaluation{cross_validation(ifolds)}=cell2mat(valP);

        evaluation_targets{cross_validation(ifolds)}=cell2mat(evaluation_targets1);
        evaluation_ids{cross_validation(ifolds)}=cell2mat(val_id);
        evaluation_indices=cell2mat(valInd);
        evaluation_indices=evaluation_indices';
        indices{1,2}='Validation_indices';
        indices{2,2}=evaluation_indices;
        Xval_indices{1,ifolds}=['Fold ' int2str(cross_validation(ifolds))];
        Xval_indices{2,ifolds}=indices;

    end

    Xval{1,1}='Training data';
    Xval{1,2}='Training targets';
    Xval{1,3}='Training ids';
    Xval{1,4}='Validation data';
    Xval{1,5}='Validation targets';
    Xval{1,6}='Validation ids';
    Xval{1,7}='Training class population';
    Xval{1,8}='Validation class population';
    Xval{1,9}='Xval indices';

```

```

Xval{1,10}='Number of folds';
Xval{2,1}=training;
Xval{2,2}=training_targets;
Xval{2,3}=training_ids;
Xval{2,4}=evaluation;
Xval{2,5}=evaluation_targets;
Xval{2,6}=evaluation_ids;
Xval{2,7}=training_class_populations;
Xval{2,8}=evaluation_class_populations;
Xval{2,9}=indices;
Xval{2,10}=cross_validation;

    reply3=input('Do you want to save the cross validation? (y/n)
','s');clc
    if reply3=='y'
        uisave('Xval')
    end
else
    reply4=input('Please load a cross validation set (Press enter):
','s');clc

    uiopen
    cd Models

    for ifold_folders=1:Xval{2,10}
        folder_fold=['Fold' int2str(ifold_folders)];
        mkdir(folder_fold)
    end
    cd ..

end

end

```

Feature Test:

```
function
[All_results_f_tst_1cmb,All_results_f_tst,reply_feature_tst]=feature_test(X
val,in_pars,classes_numbers)
[T,T_targ,rt1,V,V_targ,rt2,T_class_population,V_class_population,rt3,cross_
validation]=get_Xval_full(Xval);
show_gui=in_pars{1};
Epochs=in_pars{2};
Transfer_fn=in_pars{3};
Neurons=in_pars{4};
reply_feature_tst=input('Do you want to perform a feature test? (y/n):
','s');
clc
if reply_feature_tst=='y'
    reply_feature_n=input('Please set the number of features to be tested
(left out - max n-1): ');
    clc
    if reply_feature_n>1
        reply_make_all=input('Do you want to make models starting from one
feature? (y/n): ','s');
        else
            reply_make_all=1;
        end
        clc
        if reply_make_all=='y'
            cd Models
            for ifolds=1:length(cross_validation)
                go2current_fold=['Fold ' int2str(cross_validation(ifolds))];
                cd(go2current_fold)
                All_results_f_tst{1,cross_validation(ifolds)}=['Fold '
int2str(cross_validation(ifolds))];
                T_results2{1,1}=['Fold ' int2str(cross_validation(ifolds))];
                T_out2{1,1}=['Fold ' int2str(cross_validation(ifolds))];
                V_results2{1,1}=['Fold ' int2str(cross_validation(ifolds))];
                V_out2{1,1}=['Fold ' int2str(cross_validation(ifolds))];

                for icmbs=1:reply_feature_n
                    T_results2{2,icmbs}=['Test ' int2str(icmbs) ' feature(s)'];
                    T_out2{2,1}=['Test ' int2str(icmbs) ' feature(s)'];
                    V_results2{2,icmbs}=['Test ' int2str(icmbs) ' feature(s)'];
                    V_out2{2,1}=['Test ' int2str(icmbs) ' feature(s)'];

combos=combntrns(1:length(T{cross_validation(ifolds)}(:,1)),icmbs);
                T2=T{cross_validation(ifolds)};
                V2=V{cross_validation(ifolds)};

                folder_cmb=['Tst_' int2str(icmbs) '_feature'];
                mkdir(folder_cmb)
                cd(folder_cmb)

                [T_results2{3,icmbs},T_out2{3,icmbs},V_results2{3,icmbs},V_out2{3,icmbs}]=c
reate_NN_no_lr_mom_ftst(T2,T_targ{cross_validation(ifolds)},T_class_populat
ion,V2,V_targ{cross_validation(ifolds)},V_class_population,Neurons,Transfer
_fn,Epochs,show_gui,classes_numbers,combos);
                cd ..
            end
            cd ..

            n_comp_results{1,1}=['Fold ' int2str(ifolds)];
            n_comp_results{2,1}='Training_results';
            n_comp_results{2,2}='Training_outs';
```

```

n_comp_results{2,3}='Validation_results';
n_comp_results{2,4}='Validation_outs';

n_comp_results{3,1}=T_results2;
n_comp_results{3,2}=T_out2;
n_comp_results{3,3}=V_results2;
n_comp_results{3,4}=V_out2;
All_results_f_tst{2,cross_validation(ifolds)}=n_comp_results;
All_results_f_tst_1cmb=[];

end
reply5=input('Do you want to save the results? (y/n): ','s');
clc
if reply5=='y'
    uisave('All_results_f_tst')
end

else
    cd Models
    for ifolds=1:length(cross_validation)
        go2current_fold=['Fold' int2str(cross_validation(ifolds))];
        cd(go2current_fold)
        All_results_f_tst_1cmb{1,cross_validation(ifolds)}=['Fold '
int2str(cross_validation(ifolds))];

        T_results2{1,1}=['Fold ' int2str(cross_validation(ifolds))];
        T_results2{2,1}=['Test ' int2str(reply_feature_n) '
feature(s)'];
        T_out2{1,1}=['Fold ' int2str(cross_validation(ifolds))];
        T_out2{2,1}=['Test ' int2str(reply_feature_n) ' feature(s)'];
        V_results2{1,1}=['Fold ' int2str(cross_validation(ifolds))];
        V_results2{2,1}=['Test ' int2str(reply_feature_n) '
feature(s)'];
        V_out2{1,1}=['Fold ' int2str(cross_validation(ifolds))];
        V_out2{2,1}=['Test ' int2str(reply_feature_n) ' feature(s)'];

        combos=combntrns(1:length(T{cross_validation(ifolds)}(:,1)),reply_feature_n)
;
        T2=T{cross_validation(ifolds)};
        V2=V{cross_validation(ifolds)};

        [T_results2{3,1},T_out2{3,1},V_results2{3,1},V_out2{3,1}]=create_NN_no_lr_m
om_ftst(T2,T_targ{cross_validation(ifolds)},T_class_population,V2,V_targ{cr
oss_validation(ifolds)},V_class_population,Neurons,Transfer_fn,Epochs,show_
gui,classes_numbers,combos);

        n_comp_results{1,1}=['Fold '
int2str(cross_validation(ifolds))];
        n_comp_results{2,1}='Training_results';
        n_comp_results{2,2}='Training_outs';
        n_comp_results{2,3}='Validation_results';
        n_comp_results{2,4}='Validation_outs';

        n_comp_results{3,1}=T_results2;
        n_comp_results{3,2}=T_out2;
        n_comp_results{3,3}=V_results2;
        n_comp_results{3,4}=V_out2;

        All_results_f_tst_1cmb{2,cross_validation(ifolds)}=n_comp_results;
        All_results_f_tst=[];
        cd ..
    end
    cd ..
end

```

```

        reply5=input('Do you want to save the results? (y/n): ','s');
        clc
        if reply5=='y'
            uisave('All_results_f_tst_1cmb')
        end
    end
else
    All_results_f_tst_1cmb=[];
    All_results_f_tst=[];
end
end
end

```

Get Database:

```

function [id,data,data2,targets,classes_numbers]=Get_database
input(['Please load the database in .xls format.','...
'\n- First row must contain the titles of Ids, Targets, Feature names
and CV use (optional).','...
'\n- First column must contain the Ids.','...
'\n- Second column must contain the Study.','...
'\n- Third column must contain the Targets.','...
'\n- Fourth - n columns must contain the features.','...
'\n- n+1 - n+z column must contain the CV use (Press enter): ']);
clc

[FileName,PathName] = uigetfile('*.xls','Select the Database');
clc
disp('Loading database')
[numbers,text]=xlsread(FileName);
clc
reply_n_of_features=input('Please specify the total number of features:
');clc
features_size=reply_n_of_features;

if isempty(numbers(:,1))
    id=text(:,1);
else
    id=numbers(:,1);
end

data=numbers(:,4:3+features_size);
data=data';
targets=numbers(:,3);
classes_numbers=unique(targets);

%% SELECT INPUT FEATURES

reply_features=input('Please set the input features (All or [1,2,...,n] or
1:n): ','s');clc
if strcmp(reply_features,'All')
    reply_features=1:feature_size;
end
data2=data(str2num(reply_features),:);

end

```

Get Cross Validation Full:

```
function
[T,T_targ,T_ids,V,V_targ,V_ids,T_class_population,V_class_population,sets_i
ndices,cross_validation]=get_Xval_full(Xval)
T=Xval{2,1};
T_targ=Xval{2,2};
T_ids=Xval{2,3};
V=Xval{2,4};
V_targ=Xval{2,5};
V_ids=Xval{2,6};
T_class_population=Xval{2,7};
V_class_population=Xval{2,8};
sets_indices=Xval{2,9};
cross_validation=Xval{2,10};
end
```

Input Parameters:

```
function in_pars=input_parameters

r_show_gui=input('Do you want to show the NN interface? (y/n): ','s');clc
if r_show_gui=='y'
    show_gui=1;
else
    show_gui=0;
end
r_epochs=input('Please set the number of epochs: ');clc
Epochs=r_epochs;

stop1=1;
count1=0;
counting={'1st','2nd','3rd','th'};
while stop1
    count1=count1+1;
    if count1<=3
        r_in_lr=input(['Please set the ' counting{count1} ' input function
to be tested (ex: tansig): '], 's');clc
    else
        r_in_lr=input(['Please set the ' int2str(count1) counting{4} '
input function to be tested (ex: tansig): '], 's');clc
    end
    In_layer_TFns{count1}=r_in_lr;
    r_stop=input('Do you want to test another input function? (y/n):
','s');clc
    if r_stop=='n'
        stop1=0;
    end
end

r_n_hidd_lr=input('Please specify the number of hidden layers: ');clc

for ihidd_lrs=1:r_n_hidd_lr
    if ihidd_lrs<=3
        Layers_TFns{ihidd_lrs}=input(['Please set the function of the '
counting{ihidd_lrs} ' hidden layer (ex: purelin): '], 's');clc
    end
end
```

```

        Neurons{ihidd_lrs}=input(['Please set the number of neurons in
the ' counting{ihidd_lrs} ' hidden layer (ex: 20 or 5:5:15 or 5,20,50):
']);clc
    else
        Layers_TFns{ihidd_lrs}=input(['Please set the function of the '
int2str(ihidd_lrs) counting{4} ' hidden layer (ex: purelin): '], 's');clc
        Neurons(ihidd_lrs)=input(['Please set the number of neurons in
the' int2str(ihidd_lrs) counting{4} ' hidden layer (ex: 20 or 5:5:15 or
5,20,50): ']);clc

    end
end
Transfer_fn{1}=In_layer_TFns;
Transfer_fn{2}=Layers_TFns;

in_pars{1}=show_gui;
in_pars{2}=Epochs;
in_pars{3}=Transfer_fn;
in_pars{4}=Neurons;
end

```

Normal Run:

```

function
all_results=normal_run(Xval,in_pars,classes_numbers,reply_feature_tst)

if reply_feature_tst=='n'

[T,T_targ,yr1,V,V_targ,yr2,T_class_population,V_class_population,yr3,cross_
validation]=get_Xval_full(Xval);
    show_gui=in_pars{1};
    Epochs=in_pars{2};
    Transfer_fn=in_pars{3};
    Neurons=in_pars{4};

    cd Models
    for ifolds=1:length(cross_validation)
        T_results2{1,cross_validation(ifolds)}=['fold '
int2str(cross_validation(ifolds))];
        T_out2{1,cross_validation(ifolds)}=['fold '
int2str(cross_validation(ifolds))];
        V_results2{1,cross_validation(ifolds)}=['fold '
int2str(cross_validation(ifolds))];
        V_out2{1,cross_validation(ifolds)}=['fold '
int2str(cross_validation(ifolds))];
        go2current_fold=['Fold' int2str(cross_validation(ifolds))];
        cd(go2current_fold)

        [T_results2{2,cross_validation(ifolds)},T_out2{2,cross_validation(ifolds)},
V_results2{2,cross_validation(ifolds)},V_out2{2,cross_validation(ifolds)}]=
create_NN_no_lr_mom(T{cross_validation(ifolds)},T_targ{cross_validation(ifolds)},T_class_population,V{cross_validation(ifolds)},V_targ{cross_validation(ifolds)},V_class_population,Neurons,Transfer_fn,Epochs,show_gui,classes_numbers);
        cd ..
    end
    cd ..
    all_results{1,1}='Training results';
    all_results{1,2}='Training outputs';
    all_results{1,3}='Validation results';
    all_results{1,4}='Validation outputs';

```

```

        all_results{2,1}=T_results2;
        all_results{2,2}=T_out2;
        all_results{2,3}=V_results2;
        all_results{2,4}=V_out2;

        reply5=input('Do you want to save the results? (y/n): ','s');clc
        if reply5=='y'
            uisave('all_results')
        end
    else
        all_results=[];
    end
end
end

```