

2011

Προσομοίωση συστήματος μάζας
ελατηρίου με χρήση γενετικά
βελτιστοποιημένου νευρωνικού δικτύου

Στάθης Θεοφάνης

Technical University of Crete
10/11/2011



Τμήμα Μηχανικών Παραγωγής & Διοίκησης

Πολυτεχνείο Κρήτης

Τίτλος:

Συγγραφέας: Στάθης Θεοφάνης

Επιβλέπων καθηγητής: Σταυρουλάκης Γεώργιος

Περίληψη

Στην παρούσα εργασία αξιολογούνται οι δυνατότητες των νευρωνικών δικτύων να προσομοιωσουν συστήματα μάζας-ελατηρίου για διάφορες τιμές των παραμέτρων. Η τοπολογία των δικτύων βελτιστοποιείται με χρήση γενετικών αλγορίθμων. Τα δεδομένα εκπαίδευσης προκύπτουν από την επίλυση μη ομογενών διαφορικών εξισώσεων 2ης τάξης με χρήση του λογισμικού ODE solver.

Ευχαριστίες Στον Προτοπαπαδάκη Ευτηχη για τη βοήθεια του σε ορισμένα σημεία

Περιεχόμενα

Εισαγωγή	1
Το συστήματα μάζας-ελατηρίου	1
Λίγη φυσική.....	1
Η πειραματική διάταξη	2
Μερικά μαθηματικά.....	4
Το λογισμικό Ode solver.....	7
Νευρωνικά Δίκτυα (Neural Networks).....	10
Ιστορία.....	10
Η Δομή-Συντομη Περιγραφή.....	10
Συναρτήσεις Ενεργοποίησης.....	12
Αρχιτεκτονικές Τεχνητών Νευρωνικών Δικτύων.....	12
Εκπαίδευση-Μάθηση	13
Νευρωνικά δίκτυα χρήσεις	15
Γενετικοί Αλγόριθμοι (Genetic Algorithms)	17
Ιστορία.....	17
Με Λίγα Λόγια	18
Βασικά χαρακτηριστικά γενετικών αλγορίθμων.....	20
Πλεονεκτήματα Και Μειονεκτήματα των Γενετικών Αλγορίθμων	21
Η διάταξη που χρησιμοποιήθηκε.....	22
Παραγωγή δεδομένων εκπαίδευσης.....	22
Ο γενετικός αλγόριθμος.....	22
Αποτελέσματα	23
Επίδραση του πλήθους των κρυμμένων επιπέδων	24
Επίδραση των εποχών του γενετικού αλγορίθμου	25
Κατανομή συναρτησεων ενεργοποιησης ανα επίπεδο	25
Πληθος νευρώνων ανά επίπεδο	27
Συμπεράσματα	27
Βιβλιογραφία	27

Εισαγωγή

Το σύστημα μάζας ελατηρίου είναι από τα πιο διαδεδομένα συστήματα στον κόσμο. Έχει ένα μεγάλο πλήθος εφαρμογών τόσο σε πρακτικά προβλήματα όσο και για διδακτικό εργαλείο. Τέτοια συστήματα περιγράφονται από ένα σύνολο διαφορετικών εξισώσεων των οποίων η επίλυση δεν είναι πάντα μια εύκολη διαδικασία.

Πολλές φορές εξυπηρετεί η χρήση κάποιου προσεγγιστικού μοντέλου το οποίο περιγράφει ικανοποιητικά τη συμπεριφορά ενός υπό μελέτης συστήματος. Μια πολύ διαδεδομένη τεχνική δημιουργίας ενός προσεγγιστικού μοντέλου αποτελούν τα νευρωνικά δίκτυα.

Το συστήματα μάζας-ελατηρίου

Λίγη φυσική

Σε πολλές εφαρμογές η επιτάχυνση ενός αντικειμένου χαρακτηρίζεται από κάποιους φυσικούς νόμους όπως ο δεύτερος νόμος της κίνησης του Νιούτον $F=ma$. Συγκεκριμένα μια ωραία εφαρμογή διαφορικών εξισώσεων δεύτερης τάξης με σταθερούς συντελεστές είναι το μοντέλο ενός συστήματος ελατηρίου- μάζας. Υποθέτουμε ότι μια μάζα συνδέεται με ένα ελατήριο. Από τη φυσική, ο νόμος του Χουκ δηλώνει ότι αν ένα ελατήριο εκτοπισθεί κατά y από το σημείο ισορροπίας του, τότε η δύναμη που ασκείται από το ελατήριο είναι μια σταθερά $k>0$ πολλαπλασιαζόμενη με το ύψος της εκτόπισης του y . με άλλα λόγια

$$F_{\text{spring}} = -ky.$$

Το αρνητικό πρόσημο οφείλεται στο ότι η δύναμη θα είναι πάντα προς την αντίθετη κατεύθυνση από την εκτόπιση.

Ελατήρια χωρίς απόσβεση (χωρίς τριβή)

Μπορούμε τώρα να σχηματίσουμε ένα μοντέλο συστήματος ελατηρίου/μάζας συστήματος. Σύμφωνα με το δεύτερο νόμο του Νεύτωνα, $F=ma$ και ισχύει ότι $a=y''(t)$. Άρα παίρνουμε τη διαφορική εξίσωση δεύτερης τάξης $my''=-ky$ η οποία ισούται με $my''+ky=0$ όπου $m>0$ και $k>0$. Φυσικά μπορούμε να λύσουμε αυτό το σύστημα για κάθε τιμή των m, k εφόσον είναι ομογενής, γραμμική διαφορική εξίσωση δεύτερης τάξης. Για να μοντελοποιήσουμε αυτήν την τριβή, καταλαβαίνουμε ότι η δύναμη της τριβής είναι πάντα αντίθετη από την κατεύθυνση της κίνησης. Με άλλα λόγια, αν ένα αντικείμενο κινείται προς τα δεξιά, θα υποστεί μια δύναμη τριβής προς τα αριστερά. Επιπλέον, θα εργαστούμε με βάση την υπόθεση ότι η δύναμη λόγω της τριβής είναι ανάλογη της ταχύτητας (δηλαδή διπλασιαστεί η ταχύτητα, θα διπλασιαστεί και η δύναμη λόγω τριβής, κλπ.)

Χρησιμοποιώντας αυτές τις υποθέσεις $F_{\text{friction}} = -b\dot{y}(t)$ όπου $\dot{y}(t)$ είναι η ταχύτητα της μάζας τη στιγμή t και b είναι η σταθερά της αναλογικότητας που ονομάζεται η σταθερά της τριβής. (σημείωση ότι όταν y μετριέται σε μέτρα, t σε δευτερόλεπτα και η μάζα σε κιλά, η σταθερά b μετριέται σε $\text{Newton} \cdot \text{sec}/\text{meter}$.)

Συνδυάζοντας τη δύναμη από την τριβή και τη δύναμη από το νόμο του Χουκ, έχουμε $F_{\text{total}} = -b\dot{y} - ky$ και από τον νόμο της κίνησης του Newton $m\ddot{y} = -b\dot{y} - ky$ ή $m\ddot{y} + b\dot{y} + ky = 0$ όπου $m > 0$ είναι μάζα, $b > 0$ είναι η σταθερά της τριβής, $k > 0$ είναι το ελατήριο. (σημείωση ότι οι μονάδες δύναμης δε χρειάζεται να είναι Newton και το y δε χρειάζεται να είναι μέτρα, αλλά οι μονάδες πρέπει να είναι συμβατές, για παράδειγμα το k θα είναι μονέδες δύναμης προς μονάδες απόστασης, κλπ.) παίρνουμε γραμμική ομογενή διαφορική εξίσωση δεύτερης τάξης με σταθερούς συντελεστές. Επίσης, όπως πριν, υπάρχουν τρεις περιπτώσεις να λάβουμε υπόψιν βασιζόμενοι στις ρίζες του χαρακτηριστικού πολυωνύμου $m\ddot{y} + b\dot{y} + ky = 0$ το οποίο οδηγεί σε τρία διαφορετικά είδη συστήματος μάζας ελατηρίου.

Εφαρμογή-...

Σε αυτό το κομμάτι εισάγουμε μια εξωτερική δύναμη που επιδρά στη μάζα του ελατηρίου εκτός από τις άλλες δυνάμεις που έχουμε λάβει υπόψιν. Για παράδειγμα, υποθέτουμε ότι η μάζα ενός συστήματος μάζας/ελατηρίου ωθείται (ή έλκεται) από μια επιπρόσθετη δύναμη (ίσως το ελατήριο καθοδηγείται μηχανικά ή υπόκειται σε μαγνητικές δυνάμεις). Θα ονομάσουμε αυτήν την εξωτερική δύναμη F_{external} και υποθέσουμε ότι μεταβάλλεται με το χρόνο, δηλαδή $F_{\text{external}} = F(t)$.

Όπως πριν, $F_{\text{total}} = -b\dot{y} - ky + F_{\text{external}}$ και από το νόμο του Νεύτωνα $m\ddot{y} = -b\dot{y} - ky + F_{\text{external}}$. Άρα παίρνουμε τη μη ομογενή διαφορική εξίσωση δεύτερης τάξης (ODE): $m\ddot{y} + b\dot{y} + ky = F(t)$.

Η πειραματική διάταξη

Η πειραματική διάταξη συνίσταται από ελατήριο μάζας M του σχολικού εργαστηρίου και σώμα μάζας m , που προσδένεται στο ένα άκρο του ελατηρίου. Οι μάζα του σώματος είναι συγκρίσιμη με τη μάζα του ελατηρίου. Το άλλο άκρο του ελατηρίου στερεώνεται στο ταβάνι του εργαστηρίου. Το σύστημα ελατηρίου - σώματος τίθεται σε κατακόρυφη κίνηση (εικόνα 1). Το πείραμα επαναλαμβάνεται με σώματα διαφορετικών μαζών (m) και ελατήρια διαφορετικών σταθερών (k).

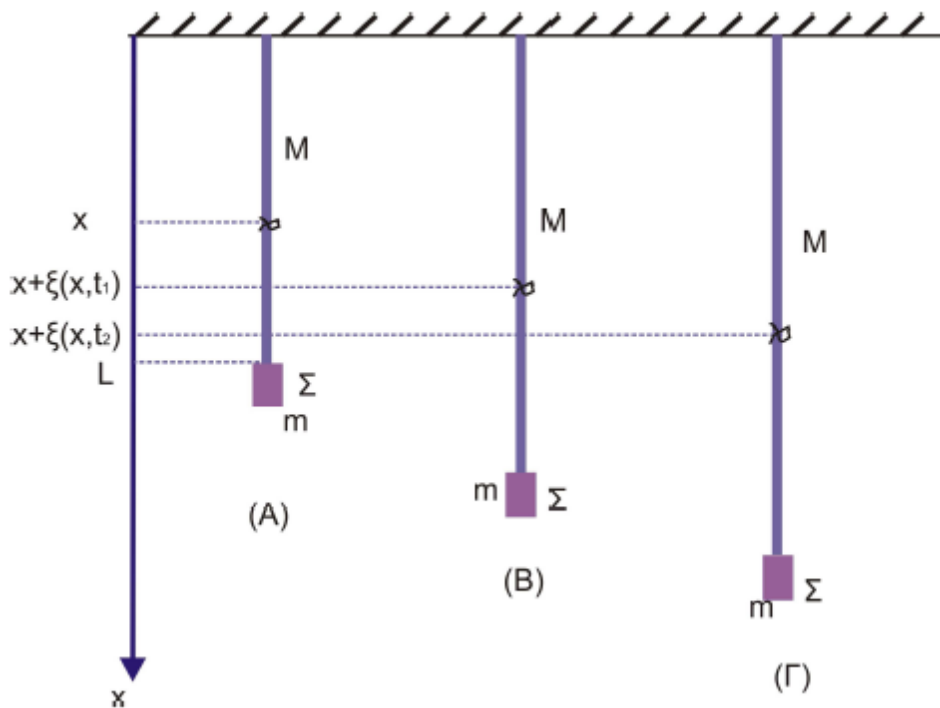
Θεωρητική ανάλυση της κίνησης του συστήματος



Έστω L το φυσικό μήκος του ελατηρίου. Εξ ορισμού, όταν το ελατήριο έχει το φυσικό του μήκος, η ελαστική δυναμική του ενέργεια είναι ίση με το μηδέν. Όταν το ελατήριο έχει το φυσικό μήκος του, οι θέσεις κάθε σπείρας προσδιορίζεται με την αλγεβρική τιμή της x , επί του κατακόρυφου άξονα Ox (εικόνα 2). Στην κατάσταση αυτή, θεωρούμε ότι η πυκνότητα σπειρών του ελατηρίου καθώς και η γραμμική πυκνότητα μάζας του είναι σταθερές:

$$n(x) = \frac{dN(x)}{dx} = \frac{N}{L} \equiv n$$

$$\rho(x) = \frac{dM(x)}{dx} = \frac{M}{L} \equiv \rho$$



Εικόνα 1 Η πειραματική διαταξη

Όταν το σύστημα ελατηρίου-σώματος κινείται κατά μήκος του άξονα Ox , το ελατήριο παραμορφώνεται κατά μήκος του ίδιου άξονα. Η σπείρα που έχει τη θέση x όταν το ελατήριο έχει το φυσικό μήκος του, μεταβάλλει τη θέση της συναρτήσει του χρόνου. Τη χρονική στιγμή t , η θέση της είναι $x + \xi(x,t)$. Η συνάρτηση $\xi(x,t)$ περιγράφει την παραμόρφωση του ελατηρίου κάθε χρονική στιγμή και προσδιορίζει τη μετατόπιση της σπείρας x τη στιγμή t . Το σώμα Σ είναι προσδεμένο στο ελεύθερο άκρο του ελατηρίου. Επομένως η κίνησή του περιγράφεται με τη συνάρτηση $\xi(L,t)$. Ως αποτέλεσμα, η μελέτη της κίνησης του συστήματος ελατηρίου-σώματος ανάγεται στην εύρεση της καταλληλούς συνάρτησης $\xi(x,t)$.

Μερικά μαθηματικά

Σύμφωνα με τις θεμελιώδεις αρχές της Κλασικής Μηχανικής(1,2,5), οι εξισώσεις κίνησης ενός μηχανικού συστήματος προκύπτουν από την ελαχιστοποίηση του συναρτησοειδούς του ολοκληρώματος της δράσης(2,3,5), που περιγράφει το μηχανικό σύστημα. Έτσι, για να βρούμε τις εξισώσεις που ικανοποιεί η συνάρτηση $\xi(x,t)$, αρκεί να εκφράσουμε αναλυτικά τη συνάρτηση Lagrange του συστήματος και να απαιτήσουμε η συνάρτηση $\xi(x,t)$ να καθιστά το ολοκλήρωμα της δράσης ακρότατο.

Το ολοκλήρωμα της δράσης του συστήματος, κατά την κίνησή του από τη χρονική στιγμή t_1 έως την t_2 , εκφράζεται από τη σχέση:

$$S[\xi(x,t)] = \int_{t_1}^{t_2} L(\dot{\xi}, \xi') dt$$

όπου $L(\xi, \xi')$ η συνάρτηση Lagrange του συστήματος και οι συντομογραφίες ξ, ξ' ορίζονται από τις σχέσεις:

$$\dot{\xi} = \frac{\partial \xi(x,t)}{\partial t} \quad \xi' = \frac{\partial \xi(x,t)}{\partial x} \quad . \quad 3.$$

Η ποσότητα ξ καθορίζει την ταχύτητα της σπείρας x , τη χρονική στιγμή t . Στην Κλασική (Νευτώνεια) Μηχανική, η συνάρτηση Lagrange υπολογίζεται από τη διαφορά της κινητικής (T) μείον τη δυναμική (V) ενέργεια του συστήματος $L=T-V$

Εφαρμόζουμε τις βασικές αρχές της Κλασικής Μηχανικής στο σύστημα που μελετάμε και καταλήγουμε στις ακόλουθες εκφράσεις για την κινητική και τη δυναμική ενέργεια του συστήματος: {Σημειώνουμε ότι στον υπολογισμό της δυναμικής ενέργειας, συνυπολογίζουμε την ελαστική δυναμική ενέργεια του ελατηρίου, τη δυναμική ενέργεια του ελατηρίου λόγω του βάρους των σπειρών του και τη δυναμική ενέργεια του σώματος Σ, λόγω του βάρους του}

$$T = \int_0^L \frac{M}{2L} (\dot{\xi}(x,t))^2 dx + \frac{1}{2} m (\dot{\xi}(L,t))^2$$

$$V = \int_0^L \left[\frac{Y}{2} (\xi'(x,t))^2 - \frac{Mg}{L} \xi(x,t) \right] dx - mg \xi(L,t)$$

5



Η σταθερά Y είναι η σταθερά ελαστικότητας του ελατηρίου. Σχετίζεται με τη δύναμη ελαστικότητας F που ασκείται στη σπείρα x , τη στιγμή t , με τη σχέση:

$$F = Y \cdot \frac{\partial \xi(x, t)}{\partial x}$$

Η σταθερά Y συνδέεται με τη σταθερά k του ελατηρίου και το φυσικό μήκος του L , με τη σχέση:

$$k = \frac{Y}{L}$$

Οι εξισώσεις κίνησης προκύπτουν από την απαίτηση:

$$\delta \int_{t_1}^{t_2} L(\xi, \xi') dt = 0$$

Χωρίς να μπούμε σε λεπτομερή ανάλυση των ενδιάμεσων βημάτων, από το συνδυασμό των (4) (5) και (7), προκύπτει ότι η συνάρτηση $\xi(x, t)$, που περιγράφει την πραγματική κίνηση του συστήματος ελατηρίου-σώματος, ικανοποιεί τις εξισώσεις:

$$\begin{aligned} \frac{M}{L} \cdot \frac{\partial^2 \xi(x, t)}{\partial t^2} - Y \cdot \frac{\partial^2 \xi(x, t)}{\partial x^2} &= \frac{Mg}{L} \\ m \cdot \frac{\partial^2 \xi(L, t)}{\partial t^2} + Y \cdot \frac{\partial \xi(x, t)}{\partial x} \Big|_{x=L} &= mg \\ \xi(0, t) &= 0 \end{aligned}$$

Διερευνούμε την ύπαρξη λύσης των εξισώσεων (8), που να αντιστοιχεί σε μια κατάσταση συμφασικής ταλάντωσης όλων των σημείων του ελατηρίου και του σώματος. Μια τέτοια λύση (αν υπάρχει) έχει τη μορφή:

$$\xi(x, t) = \xi_0(x) \cdot \sin(\omega t) + f(x)$$

Συνδυάζουμε την έκφραση (9) με τις εξισώσεις (8) και βρίσκουμε ότι η κατάσταση συμφασικής ταλάντωσης ολόκληρου του συστήματος περιγράφεται με τη συνάρτηση:

$$\xi(x, t) = A \cdot \sin(q \cdot x) \cdot \sin(\omega \cdot t) - \frac{Mg}{2YL} \cdot x^2 + \frac{(M+m)g}{Y} \cdot x$$

$$q \equiv \left(\frac{M}{YL} \right)^{1/2} \cdot \omega$$

$$\tan(qL) = \frac{Yq}{m\omega^2}$$

Διερεύνηση της λύσης (10)

1) Μπορεί κανείς να δείξει εύκολα ότι όταν η μάζα του ελατηρίου είναι αμελητέα σε σχέση με τη μάζα του σώματος Σ, από τις εξισώσεις (10) προκύπτουν οι γνωστές εξισώσεις αρμονικής ταλάντωσης με συχνότητα

$$\omega_0 = \sqrt{\frac{k}{m}}$$

2) Όταν η μάζα του ελατηρίου είναι αρκετά μικρότερη της μάζας του σώματος Σ ($M/m \ll 1$), το όρισμα της εφαπτομένης στη σχέση (10.3), είναι της τάξης $(M/m)^{1/2}$. Έτσι, μπορούμε να αναπτύξουμε την εφαπτομένη ως προς το όρισμά της σε σειρά Taylor και να την προσεγγίσουμε κρατώντας όρους μέχρι τρίτης τάξης. Στο πλαίσιο της προσέγγισης αυτής, μπορούμε να επιλύσουμε την εξίσωση (10.2) και να υπολογίσουμε τη συχνότητα ω . Προκύπτει η σχέση:

$$\omega = \omega_0 \sqrt{\left(1 - \frac{M}{3m} \right)}$$

ή

$$\omega = \sqrt{\frac{k}{m + \frac{M}{3}}}$$

Στην περίπτωση αυτή η συχνότητα ταλάντωσης του συστήματος είναι ίδια με τη συχνότητα ταλάντωσης



αβαρούς ελατηρίου ίσης σταθεράς με εκείνο της διάταξης, στο άκρο του οποίου έχει προσδεθεί σώμα μάζας ίσης με τη μάζα του σώματος Σ συν το ένα τρίτο της μάζας του ελατηρίου της διάταξης.
3) Στη γενική περίπτωση η θεμελιώδης συχνότητα του συστήματος δίνεται από τις σχέσεις (10). Το σύστημα

είναι δυνατό να πραγματοποιήσει αρμονική ταλάντωση, έτσι ώστε όλα τα σημεία του να έχουν κοινή φάση και συχνότητα, σύμφωνα με τις εξισώσεις κίνησης που περιγράφονται από σχέσεις (10).

Το λογισμικό Ode solver

Ο Ode solver της Matlab χρησιμοποιείται για να λύσει προβλήματα της μορφής :

$$\begin{aligned}\frac{dx_1}{dt} &= f_1(x_1, x_2, x_3, \dots, t) \\ \frac{dx_2}{dt} &= f_2(x_1, x_2, x_3, \dots, t) \\ \frac{dx_3}{dt} &= f_3(x_1, x_2, x_3, \dots, t) \\ &\dots\end{aligned} \quad \text{i.e.} \quad \frac{d\underline{x}}{dt} = F(t, \underline{x}).$$

Στον Ode solver της Matlab έχουμε πρόσβαση καλώντας/δημιουργώντας μια συνάρτηση της μορφής:

`[X, T] = ode**(@F, TimeSpan, Xo, Options, P1, P2, P3)` Optional

@F	Η κλήση της συναρτησης η οποία επιστρέφει ένα διανυσμα τιμων με τους ρυθμους μεταβολης
Timespan	Ένα διανυσμα γραμης το οποιο χρειαζεται
Xo	Διανυσμα αρχικών τιμών
Options (if omitted or set to [], the default settings are used)	Μια δομη δεδομενων με την οποια ο χρηστης καθοριζει διαφορες επιλογες σχετικές με τον τροπο επιλυσης
P1,P2,P3...	Επιπλέον ορισματα εισοδου για την

συναρτης F

Η F πρέπει να έχει την ακόλουθη μορφή:

```
function [dx_dt] = F(t,x,P1,P2,P3...)

dx_dt = .....

return
```

Υπάρχουν αρκετοί διαφορετικοί Ode solver που παρέχονται από την matlab.

	επιλυση	
ode 45	Ακριβης	
ode 23	Ακριβης	
ode 113	Ακριβης	
ode 15s	Πεπλεγμένη	
ode 23s	Πεπλεγμένη	
ode 23tb	Πεπλεγμένη	

Γενικά θα χρησιμοποιήσουμε τους ODE45 και ODE15 (το s υποδηλώνει ότι χρησιμοποιεί μια υπονοούμενη μέθοδο).

- Οι Ode solvers της Matlab θα είναι γενικά καλύτεροι από οτιδήποτε θα προγραμματίζε κάποιος μόνος του
- Είναι ικανοί να υπολογίσουν το σφάλμα στη λύση σε κάθε χρονικό βήμα και να αποφασίσουν πότε το χρονικό βήμα είναι πολύ μεγάλο(πολύ υψηλό σφάλμα) ή πολύ μικρό (αναποτελεσματικό).
- Είναι πιο σημαντικό να καταλάβεις αντιλήψεις όπως υπονοούμενα και σαφή σχέδια από την πραγματική λεπτομέρεια του τι κάνει κάθε ρουτίνα εσωτερικά. Ode45 (κατηγορηματικά μια μέθοδος Runge-Kutta) είναι αποτελεσματική αλλά μπορεί να γίνει ασταθής με άκαμπτα(stiff) συστήματα αυτό θα εκδηλωθεί με τον Ode solver να παίρνει όλο και μικρότερα βήματα για να αντισταθμιστεί.

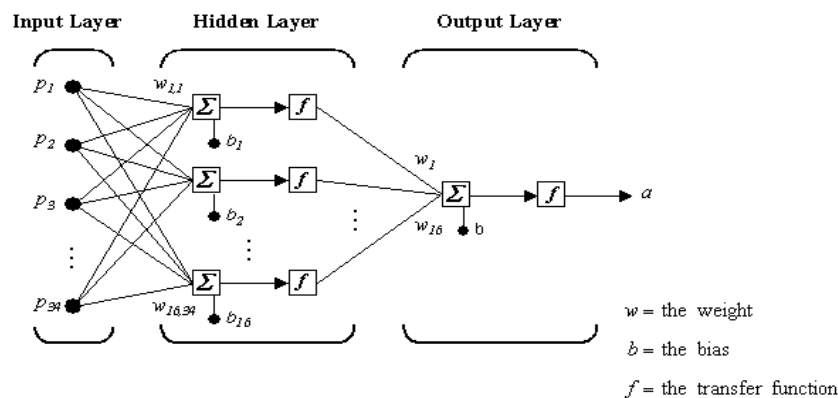


Η λύση είτε θα διαρκέσει πολύ ή το χρονικό βήμα θα μειωθεί στο σημείο όπου η ακρίβεια της μηχανής θα κάνει τη ρουτίνα να αποτύχει.

Ο ODE15 θα πρέπει να χρησιμοποιείται μόνο για άκαμπτα προβλήματα. Επειδή είναι ένα υπονοούμενο σχέδιο θα πρέπει να λύσει(μάλλον πολλά) συστήματα εξισώσεων σε κάθε χρονικό βήμα.

- Δεν χρειάστηκε να προσδιορίσουμε κάποιο μέγεθος βήματος. Ο ODE45 χρησιμοποιεί τη (σαφή) μέθοδο 4^{ης} τάξης των Runge-Kutta-Fehrberg που δίνει επίσης μια εκτίμηση του σφάλματος περικοπής σε κάθε βήμα. Ο Ode solver μπορεί να επιλέξει κάποιο μέγεθος βήματος που πληρεί την ανοχή σφάλματος που προσδιορίζουμε.
- Ode solver της Matlab θα διαφοροποιεί το μέγεθος του βήματος για να παράγουν μια λύση που είναι ακριβές σε ένα δεδομένο σφάλμα. Αυτό μπορείς να το παρακάμφθεις θέτοντας ένα μέγιστο μέγεθος βήματος (το οποίο μπορεί να αναγκάσει τον Ode solver να πάρει ένα πολύ μικρό βήμα). Γενικά είναι καλύτερο να αφήσεις τον επιλύτη να επιλέξει το μέγεθος βήματος και να χρησιμοποιήσεις τη μέθοδο τελειοποίησης. Η τελειοποίηση εξομαλύνει τη λύση παρεμβάλλοντας μεταξύ των σημείων.
- Οι επιλογές τίθενται δημιουργώντας μια δομή επιλογών με την εντολή `odeset`.
- Η επιλογή `stats` έχει ενεργοποιηθεί. Ο λύτης θα εμφανίσει στην οθόνη πληροφορίες για το πώς έχει αποδόσει.
- Όταν τρέξεις αυτόν τον κωδικό χρειάζεται σχετικά αρκετά χρόνος για να τρέξει. Οι εντολές `tic` και `toc` που περιβάλλουν τον `ode45` καλούν μέτρα και οθόνες στο χρόνο που χρειάζεται να εκτελέσεις τον `ode45`.
- Ο επιλύτης έχει αναγκαστεί να κάνει ένα πολύ μικρό βήμα για να παράγει μια πολύ μικτή λύση.
- Ο Ode solver μπορεί να κάνει πολύ γρηγορότερα βήματα (139 σε σύγκριση με τα 7557 για τον `ode45`), το οποίο μειώνει το συνολικό υπολογιστικό χρόνο.

Νευρωνικά Δίκτυα (Neural Networks)



Ιστορία

Η μελέτη υπολογιστικών συστημάτων που βασίζονται σε πρότυπα του ανθρώπινου εγκεφάλου έκανε τα πρώτα της βήματα το 1943 από τους McCulloch και Pitts οι οποίοι σχεδίασαν το πρώτο νευρωνικό δίκτυο. Η πολυπλοκότητα του ανθρώπινου εγκεφάλου είναι τέτοια έτσι ώστε απαγορεύει την πλήρη κατανόησή του. Ακόμα και η κατανόηση της λειτουργίας ενός νευρώνα του ανθρώπινου εγκεφάλου είναι φοβερά πολύπλοκη. Ο ανθρώπινος εγκέφαλος αποτελείται από 10^{10} νευρώνες, με κάθε νευρώνα να έχει αρκετές χιλιάδες συνδέσεις. Βασικά χαρακτηριστικά του ανθρώπινου εγκεφάλου είναι η αναγνώριση προτύπων (pattern recognition), ο συνειρμός, η πολυπλοκότητα και η ανεκτικότητα στο θόρυβο.

Ένας νευρώνας ενεργοποιείται όταν το σήμα εισόδου του γίνεται μεγαλύτερο από μία τιμή. Οι συνάψεις (συνδέσεις νευρώνων) μπορεί να είναι είτε διεγερτικές είτε ανασταλτικές. Ο νευρώνας έχει ένα κυτταρικό σώμα, μια δενδρική δομή εισόδων τους δενδρίτες και δενδρική δομή εξόδων τους άξονες. Οι άξονες συνδέονται με δενδρίτες άλλων νευρώνων μέσω των συνάψεων. Τα ηλεκτροχημικά σήματα εισόδων διαδίδονται από τους δενδρίτες στο κυτταρικό σώμα και έπειτα μέσω των αξόνων σε άλλους νευρώνες. Αντίστοιχες δομές ακολουθούνται και στα τεχνητά νευρωνικά δίκτυα.

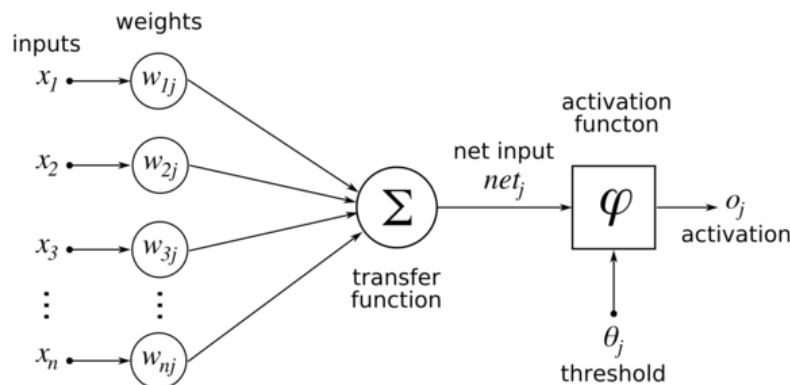
Η Δομή-Συντομη Περιγραφή

Το νευρωνικό δίκτυο είναι ένα δίκτυο από υπολογιστικούς κόμβους (νευρώνες) συνδεδεμένους μεταξύ τους. Είναι εμπνευσμένο από το Κεντρικό Νευρικό Σύστημα (ΚΝΣ), το οποίο προσπαθούν να προσομοιώσουν. Στην ουσία πρόκειται για ένα συστήματα επεξεργασίας της πληροφορίας που αποτελείται από στρώματα (layers) κάθε ένα από τα οποία αποτελείται από στοιχεία επεξεργασίας που καλούνται νευρώνες.

Οι νευρώνες είναι τα δομικά στοιχεία του δικτύου. Υπάρχουν δύο είδη νευρώνων, οι νευρώνες εισόδου και οι υπολογιστικοί νευρώνες: Οι νευρώνες εισόδου (inputs) δεν υπολογίζουν τίποτα, μεσολαβούν ανάμεσα στις εισόδους του δικτύου και τους υπολογιστικούς νευρώνες. Οι υπολογιστικοί νευρώνες πολλαπλασιάζουν τις εισόδους τους με τα



συναπτικά βάρη(weights) και υπολογίζουν το άθροισμα του γινομένου. Κάθε νευρώνας εκτελεί μία προκαθορισμένη μαθηματική λειτουργία και παράγει μία και μόνη έξοδο. Το άθροισμα που προκύπτει είναι το όρισμα της συνάρτησης ενεργοποίησης (activation function). Ο κάθε νευρώνας λειτουργεί ανεξάρτητα από τους υπόλοιπους.



Εικόνα 2 Μοντέλο νευρώνα

Το μοντέλο του νευρώνα που παρουσιάζεται παραπάνω περιλαμβάνει επίσης και ένα εξωτερικό βάρος b_k . Το βάρος b_k έχει ως αποτέλεσμα την αύξηση ή την μείωση της τιμής που δίνει σαν είσοδο το δίκτυο στην συνάρτηση ενεργοποίησης ανάλογα με το αν είναι αρνητικό ή θετικό. Το βάρος αυτό καλείται πόλωση ή κατώφλι (bias, threshold). Η τιμή της εισόδου του είναι πάντα η μονάδα, $x_0 = 1$. Με μαθηματικούς όρους, μπορούμε να περιγράψουμε έναν νευρώνα k από το παρακάτω ζεύγος εξισώσεων :

όπου $x_1, x_2, \dots, x_m$ είναι τα εισερχόμενα σήματα, $w_{k1}, w_{k2}, \dots, w_{km}$ είναι τα βάρη των συνάψεων του νευρώνα k , u_k είναι η έξοδος του γραμμικού συνδιαστή, είναι η συνάρτηση ενεργοποίησης και είναι το σήμα που δίνει σαν έξοδο ο νευρώνας k .

Η δεύτερη εξίσωση δείχνει την ιδιαίτερη σημασία του βάρους b_k . Εάν το συνολικό άθροισμα από τις υπόλοιπες εισόδους του νευρώνα είναι μεγαλύτερο από την τιμή αυτή, τότε ο νευρώνας ενεργοποιείται. Εάν είναι μικρότερο, τότε ο νευρώνας παραμένει ανενεργός.

Η δυνατότητες ενός νευρωνικού δικτύου περιγράφονται πλήρως στο ακόλουθο αποδεδειγμένο θεώρημα: Ένα νευρωνικό δίκτυο, με δύο επίπεδα ικανών για εκπαίδευση



βαρών, μπορεί να προσεγγίσει οποιαδήποτε μη γραμμική συνάρτηση (Dayhoff and DeLeo 2001).

Συναρτήσεις Ενεργοποίησης

Η συνάρτηση ενεργοποίησης ορίζει την έξοδο του νευρώνα σε σχέση με την δυνατότητα ενεργοποίησης στην είσοδό του. Υπάρχουν τρεις βασικοί τύποι συναρτήσεων ενεργοποίησης :

1. Η συνάρτηση κατώφλι (Threshold function). 
2. Τμηματικά γραμμική συνάρτηση (Piecewise Linear Function). 
3. Σιγμοειδής συνάρτηση ενεργοποίησης (sigmoid activation function). Η σιγμοειδής συνάρτηση είναι η πιο ευρέως χρησιμοποιούμενη συνάρτηση ενεργοποίησης για την κατασκευή ΤΝΔ. Ορίζεται ως μια γνησίως αύξουσα συνάρτηση που είναι ομαλή και ασυμπτωτική. 

Αρχιτεκτονικές Τεχνητών Νευρωνικών Δικτύων

Ο τρόπος με τον οποίο οι νευρώνες ενός ΤΝΔ είναι δομημένοι είναι στενά συνδεδεμένος με τον αλγόριθμο εκπαίδευσης που χρησιμοποιείται για το δίκτυο. Γενικά υπάρχουν τρεις βασικές διαφορετικές κλάσεις αρχιτεκτονικών δικτύου.

1. Πρόσω τροφοδοτούμενα Δίκτυα Ενός Επιπέδου (Single Layer Feed forward networks). Στα νευρωνικά δίκτυα οι νευρώνες του δικτύου οργανώνονται σε διάφορα επίπεδα. Η πιο απλή μορφή νευρωνικού δικτύου με επίπεδα είναι ένα νευρωνικό δίκτυο με ένα επίπεδο. Στην περίπτωση αυτή έχουμε ένα επίπεδο με κόμβους εισόδου που προβάλλονται σε ένα επίπεδο εξόδου με νευρώνες (όπου γίνονται οι υπολογισμοί), ενώ το αντίθετο δεν ισχύει, δεν μπορούμε δηλαδή να πάμε από τους νευρώνες εξόδου στους κόμβους εισόδου. Σε αυτή την περίπτωση, το δίκτυο είναι ένα αυστηρά εμπρόσθιας τροφοδότησης (feed forward) δίκτυο και καλείται Feed forward Δίκτυο Ενός Επιπέδου. Με τον όρο ένα επίπεδο εννοούμε το επίπεδο εξόδου που περιέχει και τους νευρώνες όπου γίνονται οι υπολογισμοί. Σημειώνεται ότι δεν υπολογίζεται το επίπεδο εισόδου με τους κόμβους εισόδου γιατί σε αυτό δεν γίνονται καθόλου υπολογισμοί.
2. Πρόσω τροφοδοτούμενα Δίκτυα Πολλαπλών Επιπέδων (Multilayer Feed forward Networks). Η δεύτερη κλάση εμπρόσθιας τροφοδότησης νευρωνικών δικτύων διαφέρει από την πρώτη στην ύπαρξη ενός ή περισσότερων κρυφών επιπέδων, των οποίων οι νευρώνες καλούνται κρυφοί νευρώνες. Η λειτουργία των κρυφών νευρώνων είναι να παρεμβάλλονται μεταξύ των κόμβων εισόδου και των νευρώνων εξόδου του δικτύου. Με την προσθήκη ενός ή περισσότερων κρυφών επιπέδων, το δίκτυο έχει την δυνατότητα να προσεγγίζει συναρτήσεις μεγαλύτερης πολυπλοκότητας. Οι κόμβοι εισόδου στο επίπεδο εισόδου του δικτύου παρέχουν τα στοιχεία των προτύπων



εισόδου, με την μορφή διανυσμάτων, που γίνονται είσοδοι στους νευρώνες του δεύτερου επιπέδου, δηλαδή του πρώτου κρυφού επιπέδου. Τα σήματα εξόδου των νευρώνων του δεύτερου κρυφού επιπέδου, γίνονται σήματα εισόδου στο τρίτο επίπεδο και συνεχίζεται κατά αυτό τον τρόπο η ροή των σημάτων μεταξύ των επιπέδων του δικτύου μέχρι να φτάσουν στο επίπεδο εξόδου. Συνήθως, οι νευρώνες σε κάθε επίπεδο του δικτύου, παίρνουν σαν είσοδο, τα σήματα εξόδου από τους νευρώνες του προηγούμενου επιπέδου μόνο. Το σύνολο των σημάτων εξόδου στο επίπεδο εξόδου (τελευταίο επίπεδο του δικτύου), αποτελεί και την απάντηση του δικτύου για τα δεδομένα που εισάγονται στους κόμβους εισόδου. Γενικά ένα δίκτυο με m κόμβους εισόδου, h_1 νευρώνες στο πρώτο κρυφό επίπεδο, h_2 νευρώνες στο δεύτερο κρυφό επίπεδο, ..., h_n νευρώνες στο n -οστό κρυφό επίπεδο και q νευρώνες στο επίπεδο εξόδου, αναφέρεται σαν ένα $m - h_1 - h_2 - \dots - h_n - q$ δίκτυο.

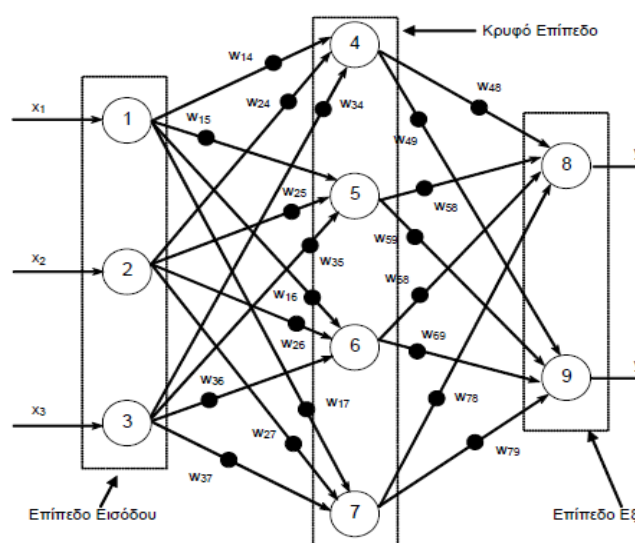
3. Αναδρομικά Τεχνητά Νευρωνικά Δίκτυα (Recurrent Neural Networks). Ένα Αναδρομικό Τεχνητό Νευρωνικό Δίκτυο (ΑΤΝΔ) διαφέρει από ένα πρόσω τροφοδοτούμενο δίκτυο στο γεγονός ότι περιέχει έναν τουλάχιστον βρόγχο ανατροφοδότησης. Αυτό σημαίνει ότι σε έναν τουλάχιστον νευρώνα, το σήμα εξόδου του επηρεάζει το σήμα που έρχεται στην είσοδο του νευρώνα. Για παράδειγμα, ένα ΑΤΝΔ μπορεί να αποτελείται από ένα μόνο επίπεδο νευρώνων όπου κάθε νευρώνας επιστρέφει το σήμα εξόδου του σαν σήμα εισόδου σε όλους τους άλλους νευρώνες του επιπέδου.

Εκπαίδευση-Μάθηση

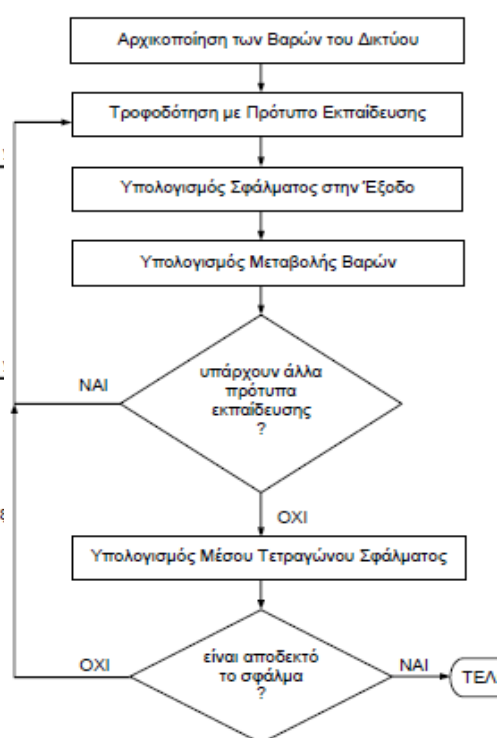
Τα βάρη αποτελούν τα στοιχεία μνήμης της αντίστοιχης διάταξης νευρώνων. Οι τιμές που παίρνουν τα βάρη καθορίζονται από την διαδικασία μάθησης. Όλες οι μέθοδοι εκμάθησης κατατάσσονται σε δύο κατηγορίες:

1. εκπαίδευση με επίβλεψη (supervised learning). Διαδικασία στην οποία απαιτείται κάποιο διάνυσμα εξόδου το οποίο αντιστοιχεί στις επιθυμητές τιμές. Οι μέθοδοι που ανήκουν σε αυτή την κατηγορία είναι η εκμάθηση με διόρθωση σφάλματος (error correction learning), η στοχαστική εκμάθηση (stochastic learning) και τα hardwired systems
2. εκπαίδευση χωρίς επίβλεψη (unsupervised learning). Διαδικασία η οποία βασίζεται μόνο στην τοπική πληροφορία κατά την διάρκεια της εκπαίδευσης. Παραδείγματα αλγορίθμων εκπαίδευσης χωρίς επίβλεψη είναι ο αλγόριθμος Hebbian, ο διαφορικός αλγόριθμος Hebbian και Min-Max αλγόριθμος.





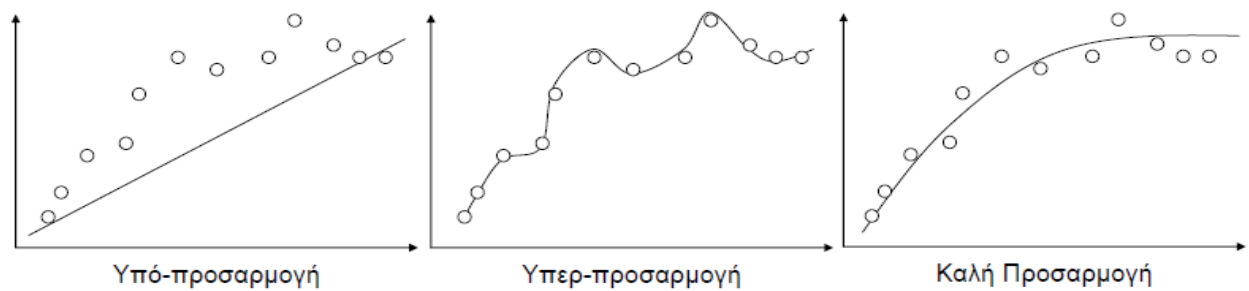
Εικόνα 3 Παράδειγμα ΤΝΔ



Εικόνα 4 Διαδικασία εκπαίδευσης ενός ΤΝΔ

Η εκπαίδευση ενός δικτύου βασίζεται στην ανάστροφη μετάδοση του σφάλματος (back propagation). Η αναπροσαρμογή των βαρών γίνεται από το επίπεδο εξόδου προς το εισόδου με εφαρμογή της διαδικασίας βελτιστοποίησης επικλινούς καθόδου (gradient descent optimization procedure) που ελαχιστοποιεί το μέσο τετραγωνικό σφάλμα E μεταξύ της εξόδου του δικτύου και της επιθυμητής εξόδου, για τα ένα συγκεκριμένο πλήθος διανυσμάτων εκπαίδευσης.

Η εκπαίδευση τελειώνει όταν τα αποτελέσματα που δίνει βρίσκονται εντός κάποιων προδιαγραφών που δίνονται από τον χρήστη ή/και την μορφή του προβλήματος. Εξαιρετικά σημαντικός παράγοντας στην εκπαίδευση ενός νευρωνικού δικτύου είναι ο αριθμός εποχών εκπαίδευσης. Ως εποχή (epoch) εννοούμε την χρήση όλων των δεδομένων εκπαίδευσης (από το δίκτυο) για μία φορά. Μικρός αριθμός εποχών εκπαίδευσης συνήθως οδηγεί σε άσχημα αποτελέσματα ενώ αντίθετα μεγάλος αριθμός εποχών εκπαίδευσης οδηγεί στην υπερεκπαίδευση και απαιτεί μεγάλο υπολογιστικό χρόνο. Υπερεκπαίδευση έχουμε όταν το νευρωνικό δίκτυο προσαρμοστεί πολύ καλά στα δεδομένα εκπαίδευσης και μόνο σε αυτά. Οι αναφορές μικρός και μεγάλος αριθμός εποχών είναι σχετικές και αναφέρονται στο εκάστοτε πρόβλημα.



Εικόνα 5 Χαρακτηριστικές περιπτώσεις εκπαιδευμένων ΤΝΔ

Νευρωνικά δίκτυα χρήσεις

Μία περίληψη των πλεονεκτημάτων και των μειονεκτημάτων των μη συμβατικών τεχνικών με έμφαση στα νευρωνικά δίκτυα περιγράφεται από τον Corr, (1994). Οι μη συμβατικές τεχνικές αποδίδουν εξίσου καλά αποτελέσματα με τις οικονομετρικές τεχνικές, λόγω της γενικευμένης λειτουργίας που επιτελούν μέσω της δυνατότητας να μιμούνται μια ευρεία ποικιλία μορφών (Curry, 2002). Ενώ λειτουργούν ως γενικευμένα μοντέλα, δεν εγγυώνται ικανοποιητικά αποτελέσματα σε προβλέψεις εκτός δεδομένων (out of sample). Παρόλα αυτά όμως, δίνουν καλύτερες αποδόσεις από τα συμβατικά μοντέλα στις περιπτώσεις όπου τα δεδομένα υποκρύπτουν μη γραμμικές σχέσεις, πράγμα το οποίο είναι το σύνηθες σε περιπτώσεις που μοντελοποιείται η απόκριση της αγοράς. Επιπλέον τα κρυφά επίπεδα των νευρωνικών δικτύων αναπτύσσουν μια εσωτερική αντιπροσώπευση των σχέσεων μεταξύ των μεταβλητών, με αποτέλεσμα να μην έχουν ανάγκη ορισμένες προϋποθέσεις που απαιτούν τα μοντέλα παλινδρόμησης, όπως είναι η έλλειψη συγγραμμικότητας μεταξύ των μεταβλητών, η ομαλότητα του θορύβου κλπ.

Επίσης τα νευρωνικά δίκτυα αποδίδουν ικανοποιητικά και σε περιπτώσεις έλλειψης κάποιων δεδομένων σε αντίθεση με τα μοντέλα παλινδρόμησης όπου δημιουργούνται σοβαρά προβλήματα (Venugopal, 1994). Ο Corr et al. (1994) συμπέραναν ότι τα νευρωνικά δίκτυα είναι κατάλληλα για πολύπλοκα φαινόμενα για τα οποία υπάρχουν ικανοποιητικά μέτρα μέτρησης της απόδοσης τους, αλλά δεν υπάρχει γνώση, για να κατανοηθούν οι σχέσεις εντός αυτών των φαινομένων. Επίσης διατύπωσαν, την άποψη ότι είναι ιδανικά για πρόβλεψη και πρόγνωση.

Από την άλλη μεριά, η κύρια ανεπάρκεια των νευρωνικών δικτύων αφορά την αιτιακή μοντελοποίηση και τον έλεγχο των υποθέσεων. Υπάρχουν, όμως, αρκετά επιτυχή αιτιακά μοντέλα που χρησιμοποιούν νευρωνικά δίκτυα. Τα οικονομετρικά μοντέλα ερμηνεύουν πολύ ευκολότερα τα αποτελέσματα και επιτρέπουν τον έλεγχο υποθέσεων. Η χρήση τεχνητών νευρωνικών δικτύων για χρηματοοικονομικές προβλέψεις συναντάται σε πλήθος ερευνών. Ενδεικτικά αναφέρονται μερικές.

Ο Andreou A. S. (2000) παρουσίασε διαφορετικές αρχιτεκτονικές νευρωνικών δικτύων, για να προβλέψει τις τιμές μετοχών στο χρηματιστήριο της Κύπρου. Επίσης, εξετάστηκε η

επίδραση οικονομικών και πολιτικών παραγόντων από το εθνικό και διεθνές περιβάλλον. Οι εφαρμογή των μοντέλων έγινε με δεδομένα του γενικού δείκτη CSE και με δεδομένα αντιπροσωπευτικών μετοχών από πέντε κλάδους της αγοράς. Τα αποτελέσματα της έρευνας προσδιόρισαν την κατάλληλη αρχιτεκτονική του νευρωνικού δικτύου, τις κατάλληλες εισόδους και τις κατάλληλες συναρτήσεις μεταφοράς. Όλες οι δοκιμές με δεδομένα, παρείχαν ικανοποιητικά αποτελέσματα πρόβλεψης.

Οι Baba N. et al. (2002) χρησιμοποίησαν ένα νευρωνικό δίκτυο και τη μέθοδο temporal Difference Learning Method, για να αναπτύξουν ένα έξυπνο σύστημα υποστήριξης αποφάσεων, ώστε να προβλέπει την τιμή των μετοχών στο χρηματιστήριο του Τόκιο ένα μήνα μπροστά. Επιλέχθηκαν 16 εταιρείες οι μετοχές των οποίων διαπραγματεύονται στο χρηματιστήριο. Δεκατέσσερις (14) μεταβλητές χρησιμοποιήθηκαν ως παράγοντες κλειδιά για την αλλαγή της τιμής κάθε μετοχής. Τα δεδομένα εκπαίδευσης αφορούν πέντε έτη. Το σύστημα παρήγαγε ικανοποιητικά κέρδη στο χρηματιστήριο του Τόκιο ακόμα και την περίοδο που το χρηματιστήριο υπέφερε από βαθιά πτώση. Η TD-Learning Method μπορεί να είναι μία σημαντική προσέγγιση για προσαρμογή στις ξαφνικές αλλαγές του χρηματιστηριακού περιβάλλοντος.

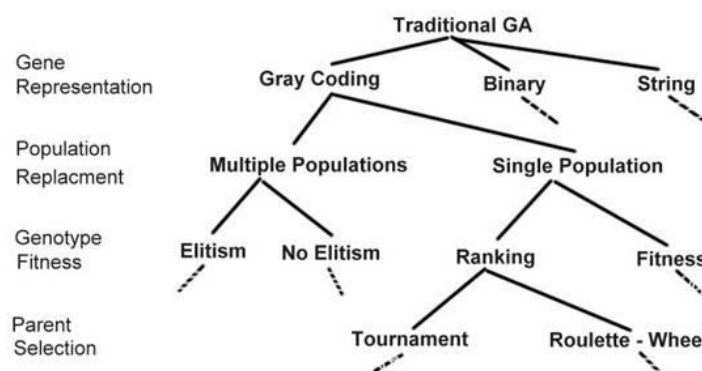
Ο Kyong. J.O. (2002) πρότεινε ένα σύστημα συναλλαγών για μετοχές βασιζόμενο σε χαοτική ανάλυση και σε μη γραμμικά μοντέλα. Αποτελείται από τέσσερις φάσεις: καθορίζει τη χρονική υστέρηση των εισόδων με τη χρήση χαοτικής ανάλυσης, προσδιορίζει σημεία αναστροφής της τάσης μιας μετοχής, προβλέπει τα σημεία αλλαγής με τη χρήση νευρωνικού δικτύου οπίσθιας τροφοδότησης και προβλέπει την τιμή της μετοχής με τη χρήση ενός νευρωνικού δικτύου οπίσθιας τροφοδότησης (BPN). Τα δεδομένα αφορούν 3069 ημερήσιες παρατηρήσεις κατά την περίοδο 1990- 2000 του δείκτη KOSPI. Για τη μέτρηση της απόδοσης του μοντέλου χρησιμοποιήθηκε το απόλυτο ποσοστό σφάλματος. Το προτεινόμενο μοντέλο δίνει σημαντικά καλύτερα αποτελέσματα σε βαθμό σημαντικότητας 1%, σε σχέση με το βασικό μοντέλο BPN.

Οι Chen A.S. et al. (2003) ανέπτυξαν ένα πιθανολογικό νευρωνικό δίκτυο (PNN), για να προβλέψουν την κατεύθυνση του δείκτη του χρηματιστηρίου της Taiwan. Το PNN έχει τρία επίπεδα σε σύγκριση με τα τέσσερα επίπεδα που έχει το βασικό PNN. Η απόδοση του συγκρίνεται με αυτή της generalized methods of moments (GMM) με φίλτρο Kalman. Οι προβλέψεις των διαφόρων στρατηγικών συναλλαγών συγκρίνονται με τη στρατηγική διακράτησης, με την τυχαία πορεία των τιμών και το παραμετρικό GMM. Το μοντέλο PNN, με οποιοδήποτε χρονικό επενδυτικό ορίζοντα, είναι ικανό να προβλέψει τη σωστή κατεύθυνση σε περισσότερες από το 50% των περιπτώσεων σε επίπεδο σημαντικότητας 5%. Το μοντέλο GMM-με φίλτρο Kalman προβλέπει σωστή κατεύθυνση πάνω από 50% σε επίπεδο σημαντικότητας 10% μόνο για πρόβλεψη έξι μήνες μπροστά. Οι κανόνες συναλλαγών του PNN υπεραποδίδουν σε σχέση με το μοντέλο GMM-Kalman, το μοντέλο τυχαίας πορείας των τιμών και τη στρατηγική αγοράς και διακράτησης. Επιπρόσθετα το PNN με πολλαπλά κατώφλια εντολών αποδίδει καλύτερα από ένα κατώφλι εντολών.



Οι Thawornwong S. et al. (2004) παρουσίασαν μία έρευνα της δυνατότητας πρόβλεψης πρόσφατων μεταβλητών στην ακρίβεια της πρόβλεψης. Το πρόσημο της πρόβλεψης χρησιμοποιείται ως απόφαση συναλλαγής. Ένας επαγωγικός αλγόριθμος δενδρικής μορφής χρησιμοποιήθηκε για να επιλεγεί το σετ των μεταβλητών. Το νευρωνικό δίκτυο που δημιούργησαν έχει ένα κρυφό επίπεδο, που χρησιμοποιεί ως συνάρτηση μεταφοράς μια σιγμοειδή υπερβολική εφαπτόμενη συνάρτηση. Τα μοντέλα πρόβλεψης συγκρίθηκαν με ένα μοντέλο γραμμικής παλινδρόμησης. Η κερδοφορία των μοντέλων χρησιμοποιήθηκε ως μέτρο σύγκρισης των μοντέλων. Τα αποτελέσματα έδειξαν ότι οι περισσότερες σωστές προβλέψεις της κατεύθυνσης της αγοράς δεν συνεπάγονται υψηλότερη κερδοφορία. Το νευρωνικό δίκτυο δίνει μεγαλύτερη κερδοφορία σε σύγκριση με τη στρατηγική αγοράς και διακράτησης.

Γενετικοί Αλγόριθμοι (Genetic Algorithms)



Ιστορία

Τα τελευταία τριάντα χρόνια, παρατηρείται ένα συνεχώς αυξανόμενο ενδιαφέρον για την ανάπτυξη συστημάτων επίλυσης προβλημάτων βασισμένων στις αρχές της φυσικής εξέλιξης. Τα συστήματα αυτού του είδους λειτουργούν διατηρώντας έναν πληθυσμό κωδικοποιημένων(συνήθως) πιθανών λύσεων του προβλήματος που προσπαθούμε να επιλύσουμε, εφαρμόζοντας πάνω σε αυτόν διάφορες διαδικασίες εμπνευσμένες από τη βιολογική εξέλιξη. Έτσι, περνώντας από γενιά σε γενιά, τα συστήματα αυτά δημιουργούν συνεχώς νέους πληθυσμούς πιθανών λύσεων εξελίσσοντας τους προηγούμενους πληθυσμούς.

Οι Γενετικοί Αλγόριθμοι (Genetic Algorithms) είναι ένα παράδειγμα τέτοιου συστήματος που μαζί με τον Εξελικτικό Προγραμματισμό (Evolutionary Programming.), τις Στρατηγικές Εξέλιξης (Evolution Strategies), τα Συστήματα Ταξινόμησης (Classifier Systems) και το Γενετικό Προγραμματισμό (Genetic Programming) αποτελούν μια κατηγορία συστημάτων επίλυσης προβλημάτων που είναι ευρύτερα γνωστή με τον όρο Εξελικτικοί Αλγόριθμοι (Evolutionary Algorithms).

Η πρώτη εμφάνιση των Γενετικών Αλγόριθμων (Γ.Α.) χρονολογείται στις αρχές του 1950, όταν διάφοροι επιστήμονες από το χώρο της βιολογίας αποφάσισαν να χρησιμοποιήσουν υπολογιστές στην προσπάθειά τους να προσομοιώσουν πολύπλοκα βιολογικά συστήματα. Η συστηματική τους ανάπτυξη όμως, που οδήγησε στην μορφή με την οποία είναι γνωστοί σήμερα, πραγματοποιήθηκε στις αρχές του 1970 από τον John Holland και τους συνεργάτες του στο Πανεπιστήμιο του Michigan. Ο ίδιος είχε αναπτύξει τον πρώτο γενετικό αλγόριθμο δεκαετία του 1960, για να δώσει τη δυνατότητα στους ηλεκτρονικούς υπολογιστές να παράγουν λύσεις σε δύσκολα προβλήματα αναζήτησης και συνδυαστικής βελτιστοποίησης, όπως η ελαχιστοποίηση συναρτήσεων και η μηχανική μάθηση.

Με Λίγα Λόγια

Οι γενετικοί αλγόριθμοι βασίζονται σε μια μίμηση της διαδικασίας εξέλιξης στην φύση, κατά την οποία νέοι και καλύτεροι πληθυσμοί μεταξύ διαφορετικών ειδών αναπτύσσονται κατά την διάρκεια της εξέλιξης. Ένας γενετικός αλγόριθμος είναι μια στοχαστική επαναληπτική διαδικασία που διατηρεί το μέγεθος του πληθυσμού (population) σταθερό σε κάθε επανάληψη η οποία ονομάζεται γενιά (generation). Η βασική λειτουργία είναι το ταίριασμα δύο ή περισσότερων λύσεων με στόχο την δημιουργία μιας καινούργιας λύσης.

Για να δημιουργηθεί μια νέα λύση εφαρμόζονται δύο τελεστές: ο τελεστής της διασταύρωσης (crossover) και ο τελεστής της μετάλλαξης (mutation). Η διασταύρωση παίρνει πλήθος ατόμων που ονομάζονται γονείς (parents) και παράγει με την ανταλλαγή τμημάτων των γονέων νέα άτομα που ονομάζονται απόγονοι (offspring). Η αξιολόγηση της κάθε λύσης γίνεται με χρήση μιας συνάρτησης καταλληλότητας (fitness function).

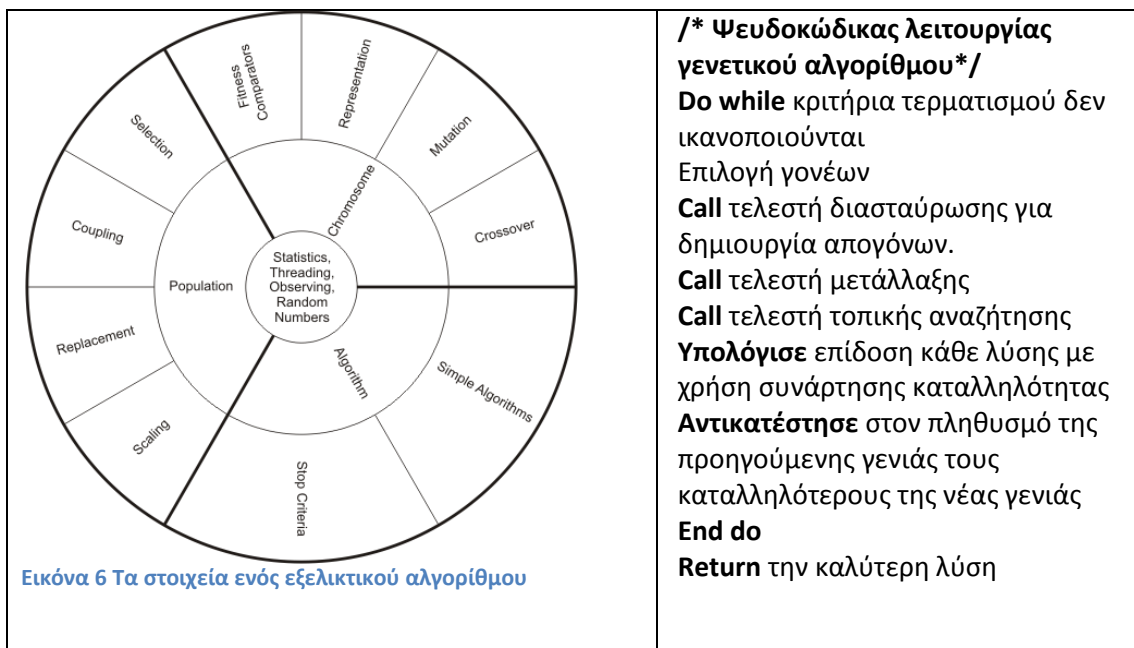
Η δομή ενός απλού Γ.Α. έχει σε γενικές γραμμές ως εξής (Michalewicz, Dasgupta et al. 1996): Κατά την διάρκεια της γενιάς t , ο Γ.Α. διατηρεί ένα πληθυσμό $P(t)$ από n πιθανές λύσεις (individuals): . Κάθε πιθανή λύση αποτιμάται και δίνει ένα μέτρο της καταλληλότητας της. Αφού ολοκληρωθεί η αποτίμηση όλων των μελών του πληθυσμού, δημιουργείται ένας νέος πληθυσμός (γενιά $t+1$) που προκύπτει από την επιλογή των πιο κατάλληλων στοιχείων του πληθυσμού της προηγούμενης γενιάς. Μερικά μέλη από τον καινούργιο αυτό πληθυσμό υφίστανται αλλαγές με την βοήθεια των γενετικών διαδικασιών της διασταύρωσης και της μετάλλαξης σχηματίζοντας νέες πιθανές λύσεις.



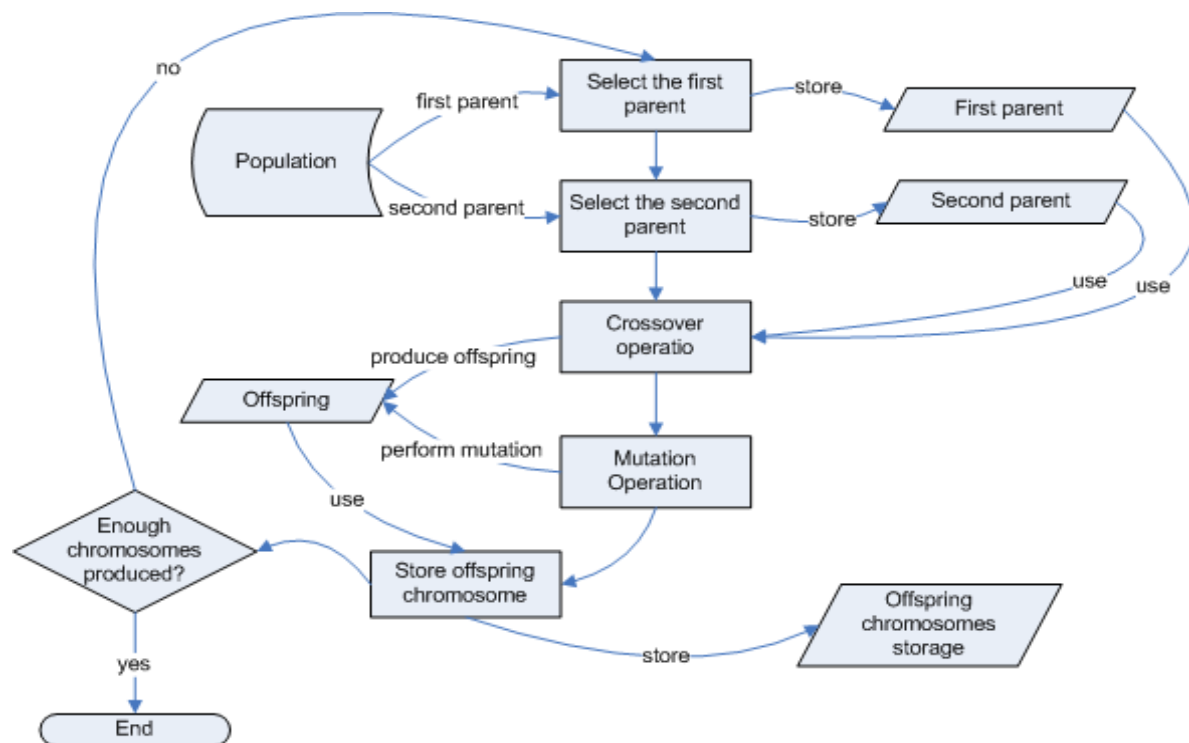
Η διασταύρωση συνδυάζει τα στοιχεία των χρωμοσωμάτων δύο γονέων για να δημιουργήσει δύο νέους απογόνους ανταλλάσσοντας τμήματα από τους γονείς. Μπορούμε να πούμε ότι η διασταύρωση εξυπηρετεί την ανταλλαγή πληροφοριών μεταξύ διαφορετικών πιθανών λύσεων. Η διαδικασία της μετάλλαξης αλλάζει αυθαίρετα ένα ή περισσότερα γονίδια ενός συγκεκριμένου χρωμοσώματος. Πραγματοποιείται με τυχαία αλλαγή γονιδίων με πιθανότητα ίση με το *ρυθμό μετάλλαξης (mutation rate)*. Η μετάλλαξη εξυπηρετεί την εισαγωγή νέων πιθανών λύσεων, διαφορετικών από τις υπάρχουσες. Με άλλα λόγια λύσεις που κανονικά δεν θα μπορούσαν να προκύψουν από τον ήδη υπάρχοντα πληθυσμό, πλέον μπορούν.

Όταν αναφερόμαστε σε Εξελικτικούς αλγόριθμους (Evolutionary Algorithms) ουσιαστικά μιλάμε για γενετικούς αλγόριθμους οι οποίοι βασίζονται στην θεωρία νέο-Δαρβινισμού¹. Ένας εξελικτικός αλγόριθμος παύει να ισχύει όταν:

1. Επιτευχθεί ο μέγιστος αριθμός γεννών (επαναλήψεων)
2. Για ένα συγκεκριμένο αριθμό γενιών δεν υπάρχει βελτίωση στην καλύτερη λύση του πληθυσμού
3. Δεν υπάρχει αλλαγή σε κανένα μέλος του πληθυσμού
4. Μια αποδεκτή λύση έχει βρεθεί



¹ Η θεωρία του νέο-Δαρβινισμού (Neo-Darwinian theory of natural evolution) συνδυάζει την θεωρία του Charles Darwin (επιβίωση του ισχυρότερου), τη θεωρία της επιλογής του August Weismann και τους νόμους διασταύρωσης του Mendel.



Εικόνα 7 Διάγραμμα ροής του τρόπου λειτουργίας ενός εξελικτικού αλγορίθμου

Βασικά χαρακτηριστικά γενετικών αλγορίθμων

Οι Γ.Α. πλεονεκτούν αισθητά στη λύση προβλημάτων αναζήτησης και βελτιστοποίησης από τις παραδοσιακές μεθόδους. Αυτό συμβαίνει, διότι διαφέρουν θεμελιωδώς από αυτές. Τα κυριότερα νέα χαρακτηριστικά που τους διαφοροποιούν, αλλά και τους δίνουν υπεροχή είναι τα εξής (Booker, Goldberg et al. 1989):

1. Οι Γ.Α. μπορούν να εφαρμοστούν με μια κωδικοποίηση του συνόλου τιμών που μπορούν να λάβουν οι μεταβλητές και όχι με τις ίδιες τις μεταβλητές του προβλήματος. Μια απλή κωδικοποίηση θα μπορούσε να γίνει θεωρώντας μια δυαδική συμβολοσειρά πεπερασμένου μήκους. Η κωδικοποίηση δεν είναι απαραίτητο να είναι πάντα δυαδική. Όπως θα φανεί και αργότερα, μπορεί να γίνει με πολλούς τρόπους, αρκετοί από τους οποίους ίσως και να μην είναι προφανείς
2. Οι Γ.Α. κάνουν αναζήτηση σε πολλά σημεία ταυτόχρονα και όχι μόνο σε ένα. Ενεργώντας ταυτόχρονα πάνω σε ένα ευρύ σύνολο λύσεων (σύνολο από συμβολοσειρές), ελαχιστοποιούν την πιθανότητα να εντοπίσουν μια ακατάλληλη κορυφή (τοπικό ακρότατο με μικρή αξία). Διαλέγοντας ένα αρχικό πληθυσμό² που να καλύπτει αντιπροσωπευτικά ένα μεγάλο εύρος τιμών μπορούν να προκύψουν ικανοποιητικά αποτελέσματα. Οι Γ.Α. χρησιμοποιούν μόνο την συνάρτηση καταλληλότητας και καμία επιπρόσθετη πληροφορία. Η μορφή των συναρτήσεων και η

² Η επιλογή κατάλληλου αρχικού πληθυσμού είναι ένα άλλο θέμα.

ύπαρξη διαφόρων περιορισμών, όπως ύπαρξη παραγώγων, συνέχεια, θόρυβος δεν αποτελούν ιδιαίτερο πρόβλημα. Ο τρόπος λειτουργίας έχει σαν αποτέλεσμα την σύγκλιση σε μια περιοχή. Εάν όμως αναζητούμε μέτωπο λύσεων (Pareto front) η πρόσθετη πληροφορία αποκτά μεγάλη σημασία και πλήθος τεχνικών για την διατήρηση της εφαρμόζονται³.

3. Οι Γ.Α. χρησιμοποιούν στοχαστικούς κανόνες μετάβασης και όχι ντετερμινιστικούς: Η χρήση στοχαστικών κανόνων μετάβασης είναι κυρίαρχο γνώρισμα των Γ.Α., χωρίς αυτό να σημαίνει ότι η όλη διαδικασία βαδίζει στην τύχη. Το στοιχείο της τύχης χρησιμοποιείται ως οδηγός για αναζήτηση σε περιοχές που αναμένεται να δώσουν καλά αποτελέσματα.

Πλεονεκτήματα Και Μειονεκτήματα των Γενετικών Αλγορίθμων

Οι Γ.Α. εφαρμόζονται σε πάρα πολλά πεδία. Αυτό οφείλεται στην ελευθερία επιλογής των κριτηρίων που καθορίζουν την επιλογή μέσα στο τεχνικό περιβάλλον. Ένας Γ.Α. είναι μια τεχνική που χρειάζεται έναν αρχικό πληθυσμό, μια συνάρτηση καταλληλότητας και μερικούς τελεστές. Αυτός είναι ο λόγος που τους επιτρέπει να χρησιμοποιηθούν στην οικονομία, στο σχεδιασμό μηχανών, στην επίλυση μαθηματικών εξισώσεων, στην εκπαίδευση Νευρωνικών Δικτύων και σε πολλούς άλλους τομείς. Η δομή των Γ.Α. οδηγεί σε σημαντικά πλεονεκτήματα αλλά και μειονεκτήματα τα οποία αναλύονται παρακάτω.

Πλεονεκτήματα

1. Είναι εύκολα επεκτάσιμοι και εξελίξιμοι. Μπορούν εύκολα να συνεργαστούν με τα υπάρχοντα μοντέλα και συστήματα, με προσθετικό τρόπο, μη απαιτώντας την επανασχεδίαση τους. Μπορούν εύκολα να συνεργαστούν με τον υπάρχοντα κώδικα και να χρησιμοποιηθούν σε συνδυασμό με άλλες μεθόδους, χωρίς μεγάλο κόπο. Αυτό συμβαίνει, διότι χρησιμοποιούν μόνο πληροφορίες της διαδικασίας ή συνάρτησης που πρόκειται να βελτιστοποιήσουν, δίχως να ενδιαφέρει άμεσα ο ρόλος της μέσα στο σύστημα ή η όλη δομή του συστήματος.
2. Δεν ενδιαφέρει η σημασία της υπό εξέταση πληροφορίας. Η μόνη "επικοινωνία" του Γ.Α. με το περιβάλλον του είναι η αντικειμενική συνάρτηση. Αυτό εγγυάται την επιτυχία του ανεξάρτητα από την σημασία του προβλήματος. Βέβαια, δεν σημαίνει ότι δεν υπάρχουν άλυτα προβλήματα για τους Γ.Α. Όπου όμως δεν τα καταφέρνουν, η αιτία είναι η φύση του χώρου που ερευνούν και όχι το πληροφοριακό περιεχόμενο του προβλήματος.
3. Είναι μία μέθοδος που κάνει ταυτόχρονα εξερεύνηση του χώρου αναζήτησης και εκμετάλλευση της ήδη επεξεργασμένης πληροφορίας. Ο συνδυασμός αυτός σπάνια συναντάται σε οποιαδήποτε άλλη μέθοδο. Με το τυχαίο ψάξιμο γίνεται καλή εξερεύνηση του χώρου, αλλά δεν γίνεται εκμετάλλευση της πληροφορίας. Αντίθετα, με το hill+-climbing γίνεται καλή εκμετάλλευση της πληροφορίας, αλλά όχι καλή εξερεύνηση.

³ Η σημασία της πρόσθετης πληροφορία γίνεται αντιληπτή στον τομέα της Πολυκριτήριας εξελικτικής βελτιστοποίησης (Evolutionary Multicriteria Optimization)

Συνήθως τα δύο αυτά χαρακτηριστικά είναι ανταγωνιστικά και το επιθυμητό είναι να συνυπάρχουν και τα δύο προς όφελος της όλης διαδικασίας. Οι Γ.Α. επιτυγχάνουν το βέλτιστο συνδυασμό εξερεύνησης και εκμετάλλευσης, πράγμα που τους κάνει ιδιαίτερα αποδοτικούς και ελκυστικούς.

4. Έχουν από τη φύση τους τα στοιχεία του παραλληλισμού και επιδέχονται παράλληλη υλοποίηση. Παρ' όλα αυτά, επειδή οι Γ.Α. χρησιμοποιούν διασταύρωση, η οποία αποτελεί ένα τρόπο διανομής της πληροφορίας, υπάρχουν δύο εξαιρετικές παραλλαγές των ΓΑ που τους επιτρέπουν πλήρη παραλληλία. Η πρώτη μέθοδος, η οποία είναι και η πιο άμεση, είναι να υπάρχει ένας γενικός πληθυσμός με πολλούς επεξεργαστές για την αποτίμηση των λύσεων των ατόμων του πληθυσμού. Η δεύτερη μέθοδος, η οποία συχνά αποκαλείται το απομονωμένο μοντέλο (island model), διατηρεί ξεχωριστούς υπο-πληθυσμούς. Κάθε τόσο ένα άτομο από έναν από τους υπο-πληθυσμούς επιτρέπεται να μεταναστεύσει σε κάποιον άλλο υπο-πληθυσμό. Με αυτό τον τρόπο γίνεται η διανομή της πληροφορίας ανάμεσα στους υπο-πληθυσμούς.

Μειονεκτήματα

1. Προβλήματα εξοικείωσης με τη γενετική. Η προσπάθεια αντιγραφής των βιολογικών μεθόδων σε ένα αλγόριθμο παρουσιάζει πολλά προβλήματα. Η ελλιπής γνώση διαφόρων αρχών της γενετικής μπορεί να οδηγήσει σε υλοποίηση λιγότερο αποτελεσματικών αλγορίθμων. Σε πολλές εφαρμογές, έχουν αναφερθεί λειτουργίες των Γ.Α. που δεν είναι δανεισμένες από τη φύση ή που έχουν υποστεί σημαντικές αλλαγές.
2. Το πρόβλημα του χρόνου και της σύγκλισης. Δεν υπάρχει μαθηματικό υπόβαθρο που να αποδεικνύει την σύγκλιση του αλγορίθμου σε ολικό ακρότατο. Ως εκ τούτου, δεν υπάρχει τρόπος υπολογισμού του χρόνου εκτέλεσης που απαιτεί ο αλγόριθμος. Οι όποιες προσπάθειες, μέχρι στιγμής, βασίζονται κυρίως σε μοντέλα Markov. Στην πράξη οι Γ.Α. έχουν δείξει ότι να αντιμετωπιστούν καλύτερα προβλήματα που έχουν πολλές, δύσκολα προσδιορισμένες λύσεις από τις συμβατικές τεχνικές.

Η διάταξη που χρησιμοποιήθηκε

Παραγωγή δεδομένων εκπαίδευσης

Ο γενετικός αλγόριθμος

Πληροφορίες για το γενετικό αλγόριθμο που χρησιμοποιήσα

Κάθε άτομο του πληθυσμού αποτελείται από:



- Τον αριθμό εποχών εκπαίδευσης
- Τον αριθμό κρυμμένων επιπέδων
- Τον αριθμό νευρώνων ανά κρυμμένο επίπεδο
- Τις συναρτήσεις ενεργοποίησης ανά επίπεδο
- Την καταλληλότητα του ατόμου (με άλλα λόγια ένα μέτρο εκτίμησης της απόδοσης)
- Το νευρωνικό δίκτυο που δημιουργείται με χρήση των παραπάνω παραμέτρων με χρήση της συνάρτησης `new ff`

Αρχικά έχουμε τρεις ομάδες από πιθανές λύσεις (νησιά) με οκτώ άτομα σε κάθε ομάδα. Ο αριθμός εποχών του γενετικού αλγόριθμου ήταν ίσος με οκτώ.

Το εύρος των τιμών στα κρυμμένα επίπεδα δίνεται από το χρήστη και ήταν ίσο από 1 έως 3. Αντίστοιχα το πλήθος νευρώνων ανά επίπεδο κυμαίνεται από 8 έως 11. Συνολικά υπάρχουν 8 διαφορετικές συναρτήσεις που χρησιμοποιήθηκαν.

Η συνάρτηση καταλληλότητας μετράει το μέσο τετραγωνικό σφάλμα των αποτελεσμάτων του κάθε ατόμου. Το άτομο με το μικρότερο μέσο τετραγωνικό σφάλμα επιλέγεται πάντα από γενιά σε γενιά. Κατά τη φάση σύγκρισης των λύσεων χρησιμοποιούνται δυο άτομα από τον υπάρχοντα πληθυσμό καθώς και δυο απόγονοι αυτών. Στην επόμενη φάση θα περάσουν οι δυο καλύτερες λύσεις.

Η ρουτίνα μετάλλαξης επιτρέπει την αλλαγή ενός ή και περισσότερων παραμέτρων κάποιου ατόμου. Οι νέες τιμές που θα προκύψουν μπορεί να βρίσκονται και εκτός ορίων που έχει θέσει ο χρήστης. Ο αλγόριθμος χρησιμοποιεί ανταλλαγή ατόμων μεταξύ των ομάδων (νησιών). Ανά τέσσερις εποχές επιλέγονται τυχαία κάποια άτομα μεταξύ δυο νησιών και συγκρίνονται μεταξύ τους. Αν κάποιο άτομο του δεύτερου νησιού προκύψει καλύτερο από αυτό του πρώτου με το οποίο συγκρίνεται ανταλλάσσουν θέση. Ο σκοπός αυτής της διαδικασίας είναι να μαζέψει τις καλύτερες λύσεις στην πρώτη ομάδα.

Για να ενεργοποιηθεί η απόδοση του κάθε ατόμου χρησιμοποιούνται δεδομένα που έχουν προκύψει από τον `ode solver` του λογισμικού... συγκρίνοντας τα αποτελέσματα που δίνει το νευρωνικό δίκτυο με τα αντίστοιχα του `ode solver` υπολογίζεται το μέσο τετραγωνικό σφάλμα και αξιολογούνται οι λύσεις (άτομα). Στο τέλος του αλγόριθμου επιστρέφεται η καλύτερη δυνατή λύση μεταξύ όλων.

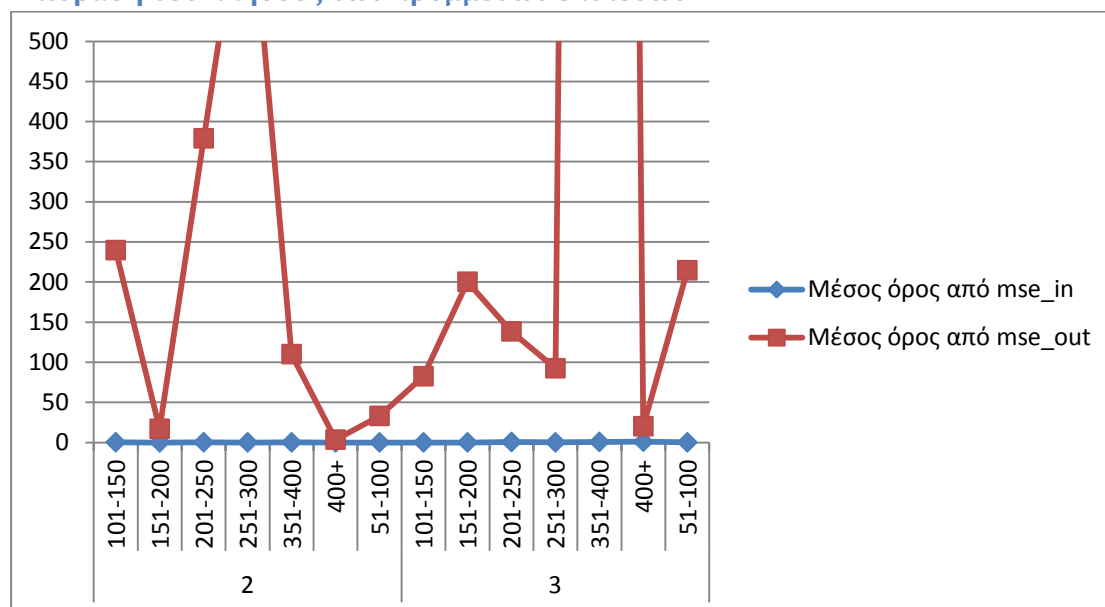
Αποτελέσματα



Το νευρωνικό δίκτυο παράγει δυο εξόδους. Η πρώτη αφορά την μετατόπιση του ελατηρίου και η δεύτερη την ταχύτητα. Η διαδικασία επαναλαμβάνεται για ένα πλήθος συντελεστών c . σε όλες τις περιπτώσεις το νευρωνικό δίκτυο προβλέπει ακριβώς τις τιμές μετατόπισης και ταχύτητας. Στη δε ταχύτητα παρουσιάζεται μια καθυστέρηση φάσης.

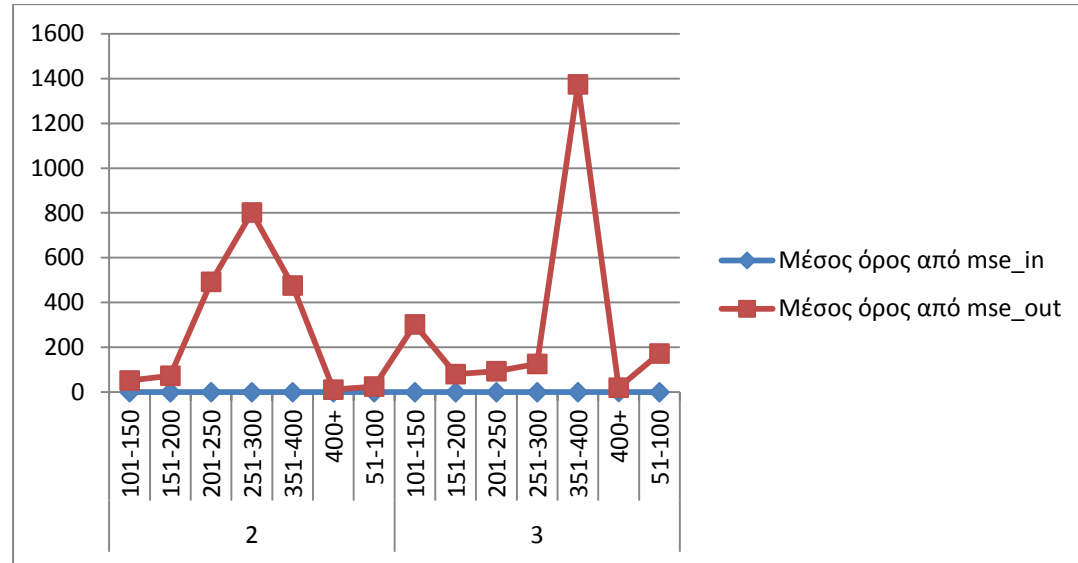
Η χρήση δύο ανεξάρτητων νευρωνικών δικτύων ένα για να προβλέπει μετατόπιση και ένα για να προβλέπει ταχύτητα δεν εφαρμόστηκε σε .?. στο σημαντικό υπολογιστικό χρόνο που θα απαιτούσε. Επιπλέον το νευρωνικό δίκτυο αποδίδει πολύ καλά ακόμα και όταν παρουσιάζεται θόρυβος στα δεδομένα και πιο συγκεκριμένα στην ασκούμενη δύναμη $f(f)$

Επίδραση του πλήθους των κρυμμένων επιπέδων



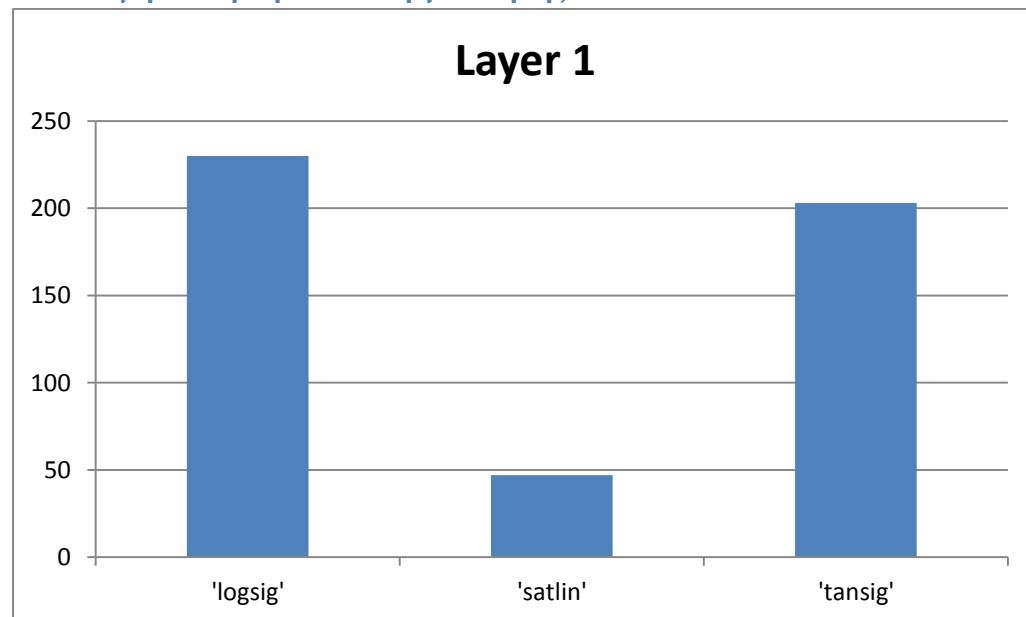
Εικόνα 8

Επίδραση των εποχών του γενετικού αλγορίθμου



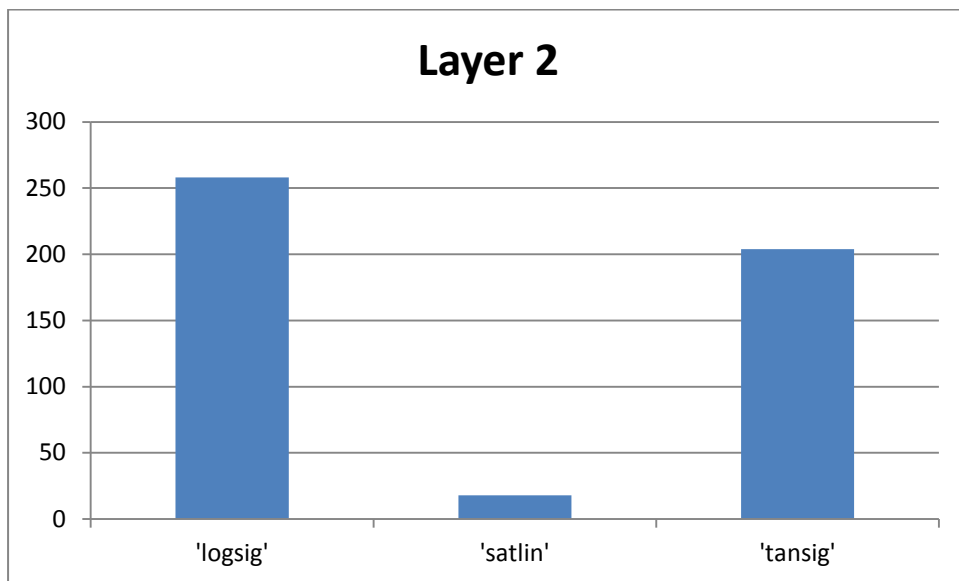
Εικόνα 9 Σύγκριση μέσου όρου σφάλματος εντός και εκτός δείγματος

Κατανομή συναρτήσεων ενεργοποίησης ανά επίπεδο

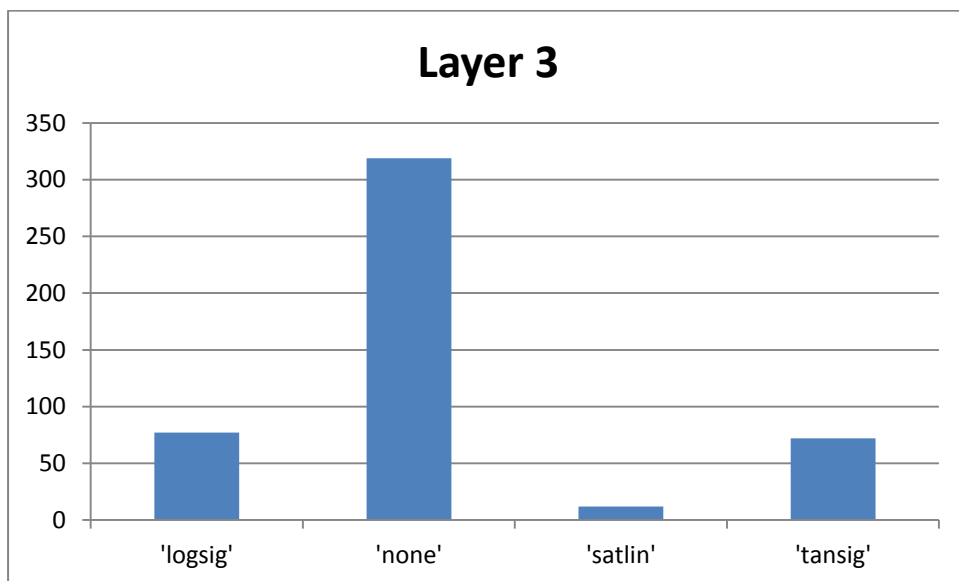


Εικόνα 10 Η κατανομή της satlin είναι αρκετά μικρότερη των άλλων δύο για 1 επίπεδο



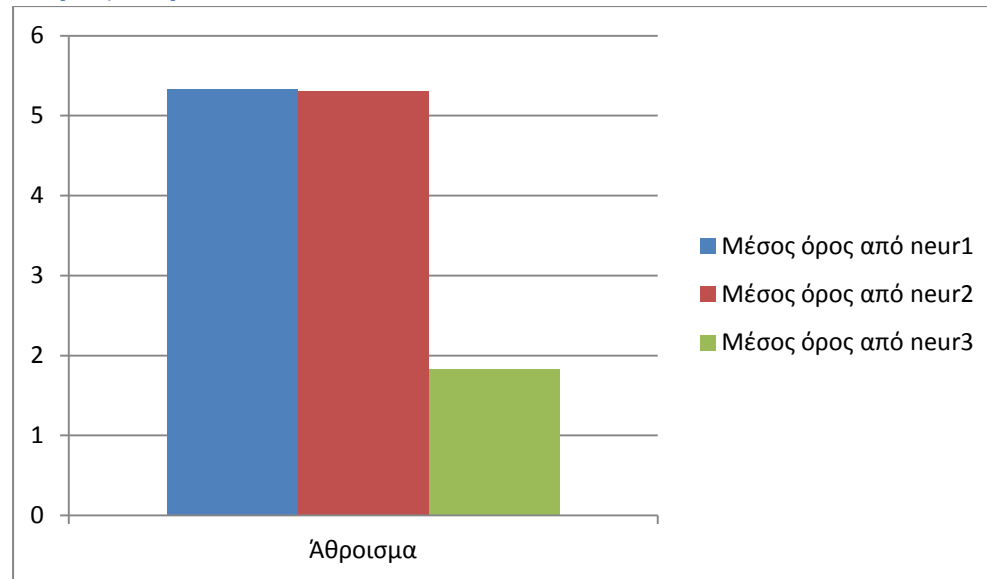


Εικόνα 11 Ομοίως και για 2 επίπεδα (layers)



Εικόνα 12 Το μεγαλύτερο πλήθος δεν έχει 3 επίπεδα όπως προέκυψε. Όμως και από αυτά που έχουν πάλι θα προτιμηθούν οι logsig και tansig έναντι της satlin

Πλήθος νευρώνων ανά επίπεδο



Εικόνα 13 Λίγα επίπεδα έχουν 3 νευρώνες. Συνήθως έχουν 1 ή 2

Συμπεράσματα

Τα νευρωνικά δίκτυα παρουσιάζουν πολύ καλή προσαρμογή στα δεδομένα. Η χρήση του γενετικού αλγόριθμου βοηθάει στην εύρεση της καταλληλότερης τοπολογίας ανά περίπτωση. Παρόλο που η τοπολογία διαφέρει κατά περίπτωση, σε ένα μεγάλο πλήθος των εξεταζόμενων περιπτώσεων παρατηρούνται κάποια κοινά αποτελέσματα.

Ο αριθμός κρυμμένων επιπέδων συγκλίνει στην τιμή 2.

Βιβλιογραφία

Agard, B. and Kusiak, A. (2004), Data mining based methodology for the design of product families, International Journal of Production Research, 42(15), 2955-2969.



- Amaratunga, D. and Cabrera, J. (2004), *Exploration and Analysis of DNA Microarray and Protein Array Data*, Wiley-Interscience. 284
- Armstrong, N. and van de Wiel, M. (2004), Microarray data analysis: From hypotheses to conclusions using gene expression data, *Cellular Oncology*, 26(5/6), 279-290.
- Bäck, T., Fogel DB, Michalewicz Z., and Beck T. (2000a), *Evolutionary Computation 1: Basic Algorithms and Operators*, Institute of Physics Publishing, Bristol, UK.
- Bäck, T., Fogel DB, Michalewicz Z., and Beck T. (2000b), *Evolutionary Computation 2: Advanced Algorithms and Operators*, Institute of Physics Publishing, Bristol, UK.
- Baldi, P. and Hatfield, GW (2002), *DNA Microarrays and Gene Expression: From Experiments to Data Analysis and Modeling*, Cambridge University Press, Cambridge, UK.
- Bar-Or, A., Keren, D., Schuster, A. and Wolff, R. (2005), Hierarchical Decision Tree Induction in Distributed Genomic Databases, *IEEE Transactions on Knowledge & Data Engineering*, 17(8), 1138-1151.
- Bergeron, B. (2003), *Bioinformatics Computing*, Prentice-Hall, Upper Saddle River, NJ.
- Bourbakis, N. and Karypis, G. (2005), Preface in bioinformatics, *International Journal of Artificial Intelligence Tools*, 14(4), 559-560.
- Brown, SM (2000), *A Biologist's Guide to Biocomputing and the Internet*, Eaton Publishing, Natick, MA
- Claverie, JM and Notredame, C. (2003) *Bioinformatics for Dummies*, Wiley Publishing Inc., NY.
- Coello, CA, Van Veldhuizen DA, and Lamont, GB (2002), *Evolutionary Algorithms for Solving MultiObjective Problems*, Plenum Press, US.
- Coppin, B. (2004), *Artificial Intelligence Illuminated*, Jones and Bartlett Publishers.
- Deb, K. and Kalyanmoy D. (2001), *Multi-Objective Optimization Using Evolutionary Algorithms*, John Wiley & Sons, Inc.
- Draghici, S. (2003), *Data Analysis Tools for DNA Microarrays*, Chapman & Hall/CRC, Boca Raton, FL.
- Eiben, AE and Smith, JE (2003), *Introduction to Evolutionary Computing*, Springer-Verlag, Berlin, Germany.
- GeneSight , Bioscovery, El Segundo, CA 90245, www.biodiscovery.com/genesight.asp , viewed July 15, 2005.



- Ghosh, A. and Nath B. (2004), "Multi-objective rule mining using genetic algorithms", *Information Science: An International Journal*, 163(1-3), 123-133.
- Giráldez, R., Aguilar-Ruiz, J. and Riquelme, J. (2005), Knowledge-Based Fast Evaluation for Evolutionary Learning, *IEEE Transactions on Systems, Man & Cybernetics: Part C*, 35(2), 254-261.
- Guan, S. and Zhu, F. (2005), An Incremental Approach to Genetic-Algorithms-Based Classification, *IEEE Transactions on Systems, Man & Cybernetics: Part B*, 35(2), 227-239.
- Hardiman, G (2003), *Microarrays Methods and Applications: Nuts & Bolts*, DNA Press, www.dnapress.net.
- Hoppe, C. (2005), Bioinformatics: Computers or clinicians for complex disease risk assessment? *European Journal of Human Genetics*, 13(8), 893-894.
- Krawetz, SA and Womble, DD (2003), *Introduction to Bioinformatics*, Humana Press, Totowa, NJ.
- Li, R. and Wang, Z. (2004), Mining classification rules using rough sets and neural networks, *European Journal of Operational Research*, 157, 439-448. 285
- Lindlöf, A., Lubovac, Z. and Michael, H. (2005), Simulations of Simple Artificial Genetic Networks Reveal Features in the Use of Relevance Networks, *Silico Biology*, 5(3), 239-249.
- McLachlan, GJ, Do KA, and Ambroise C (2004), *Analyzing Microarray Gene Expression Data*, WileyInterscience.
- Mitchell, M. (1999), *An Introduction to Genetic Algorithms*, The MIT Press, Cambridge, MA.
- NeuralWare (2003), *NeuralWorks Predict® Getting Started Guide for Windows*, Carnegie, PA.
- NeuralWare (2005), *NeuralWare Predict*, <http://www.neuralware.com/products.jsp>, viewed July 1, 2005.
- Parmigiani, G., Garrett, ES, Irizarry, RA, and Zeger, SL, Editors (2003), *The Analysis of Gene Expression Data: Methods and Software*, Springer-Verlag, Inc., New York.
- Pevsner, J (2003), *Bioinformatics and Functional Genomics*. Wiley-Liss. 551-562.
- Schena, M (2003), *Microarray Analysis*, Wiley-Liss.
- Segall, RS and Zhang, Q (2004), *Applications of Modern Heuristics and Data Mining Techniques in Knowledge Discovery*, Proposal submitted to the Summer Research Grant Committee, College of Business, Arkansas State University, State University.



Singh, GB (2003a), “Statistical Modeling of DNA Sequences and Patterns,” Chapter 22 in Krawetz, SA and Womble, DD (2003), Introduction to Bioinformatics, Humana Press, Totowa, NJ.

Singh, GB (2003b), “Statistical Mining of the Matrix Attachment Regions (MARs) in Genomic Sequences, Chapter 23 in Krawetz, SA and Womble, DD (2003), Introduction to Bioinformatics, Humana Press, Totowa, NJ.

Soransen, K. and Janssens, G. (2003), Data mining with genetic algorithms on binary trees, European Journal of Operational Research, 151, 253-264.

Speed, T. (2003), Statistical Analysis of Gene Expression Microarray Data, Chapman & Hall/CRC.

Stekel, D (2004), Microarray Bioinformatics, Cambridge University Press, New York.

Winker, P. and Gilli, M. (2004), Applications of optimization heuristics to estimation and modelling problems, Computational Statistics & Data Analysis, 47(2), 211-223.

Wit, E and McClure, J (2005), Statistics for Microarrays: Design, Analysis and Inference, John Wiley & Sons.

