



## ΔΙΑΦΟΡΟΕΞΕΛΙΚΤΙΚΑ ΠΙΘΑΝΟΤΙΚΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

**Β.Α. Γεωργίου, Ν.Γ. Παυλίδης  
Φ.Δ. Αλεβίζος, Μ.Ν. Βραχάτης**

Τμήμα Μαθηματικών

*University of Patras Artificial Intelligence Research Center (UPAIRC)*

*Πανεπιστήμιο Πατρών, 26110 Πάτρα,*

*Email: {vlg, nrvan, philipos, vrahatis}@math.upatrs.gr*

### ΠΕΡΙΛΗΨΗ

Στην εργασία αυτή προτείνονται τα Διαφοροεξελικτικά Πιθανοτικά Νευρωνικά Δίκτυα (Differential Evolution Probabilistic Neural Networks). Το κλασικό πιθανοτικό νευρωνικό δίκτυο είναι ένα μοντέλο που χρησιμοποιείται για ταξινόμηση και αναγνώριση προτύπων σε προκαθορισμένες κλάσεις και είναι επίσης γνωστό σαν Διαχωριστική Ανάλυση με χρήση πυρήνων (Kernel Discriminant Analysis). Το μοντέλο αυτό ενσωματώνει τον κανόνα ταξινόμησης κατά Bayes, δηλαδή ταξινομεί τις νέες παρατηρήσεις σύμφωνα με τις εκ των υστέρων πιθανότητες τους.

Τα μοντέλα που προτείνονται ενσωματώνουν διαφοροεξελικτικούς αλγόριθμους με σκοπό τη βελτιστοποίηση των παραμέτρων λείανσης (smoothing) των πυρήνων (kernels) του πιθανοτικού νευρωνικού δικτύου. Οι διαφοροεξελικτικοί αλγόριθμοι βελτιστοποίησης είναι αλγόριθμοι που βασίζονται σε πληθυσμούς (population-based) και δεν απαιτούν υπολογισμούς παραγώγων παρά μόνο συναρτησιακούς υπολογισμούς. Τα προτεινόμενα μοντέλα εφαρμόζονται σε τρία προβλήματα του πραγματικού κόσμου (real-world applications) από το χώρο της βιοϊατρικής και συγκρίνονται με τεχνητά νευρωνικά δίκτυα εμπρόσθιας τροφοδότησης με ενθαρρυντικά αποτελέσματα.

## 1. ΕΙΣΑΓΩΓΗ

Η ταξινόμηση και η αναγνώριση προτύπων είναι ένα αντικείμενο που μπορεί να εκτελεστεί με μία μεγάλη πληθώρα μοντέλων. Μια ενδιαφέρουσα κατηγορία προβλημάτων στην οποία χρησιμοποιούνται νευρωνικά δίκτυα καθώς και άλλα σύγχρονα μοντέλα είναι προβλήματα βιοϊατρικής. Συγκεκριμένα, προβλήματα όπως η διάγνωση καλοήθους ή κακοήθους καρκίνου, η διάγνωση διαβήτη ή ασθενειών καρδιάς είναι από τα πλέον σημαντικά τα οποία χρήζουν ιδιαίτερης προσοχής καθώς είναι άρρηκτα συνδεδεμένα με τον άνθρωπο. Η χρήση πιθανοτικών νευρωνικών δικτύων σε τέτοια προβλήματα έδωσε καλά αποτελέσματα [ Georgiou *et al* (2006), Georgiou *et al* (2004A), Georgiou *et al* (2004B), Huang (2002), Orr (1997), Wang *et al* (1998)].

Τα Πιθανοτικά Νευρωνικά Δίκτυα (ΠΝΔ) αναπτύχθηκαν από τον Specht το 1990 [Specht (1990)]. Αποτελούν μια κλάση νευρωνικών δικτύων που συνδυάζει κάποια από τα καλύτερα χαρακτηριστικά της στατιστικής αναγνώρισης προτύπων και των νευρωνικών δικτύων εμπρόσθιας τροφοδότησης (feedforward neural networks, FNNs). Τα ΠΝΔ είναι η υλοποίηση της διαχωριστικής ανάλυσης πυρήνων (kernel discriminant analysis) με βάση τα νευρωνικά δίκτυα. Το κλασικό ΠΝΔ μπορεί να χαρακτηριστεί σαν μια ευφυής μνήμη διότι κάθε πρότυπο εκπαίδευσης (training pattern) αποθηκεύεται σε ένα νευρώνα του δικτύου [Berthold *et al* (1998)]. Χαρακτηριστικές ιδιότητες των ΠΝΔ είναι ο μικρός χρόνος εκπαίδευσης και η παραγωγή αποτελεσμάτων με χρήση των εκ των υστέρων (posterior) πιθανοτήτων κατά Bayes. Αυτές οι ιδιότητες έχουν σαν αποτέλεσμα τις μεγάλες απαιτήσεις σε μνήμη και την αργή ταχύτητα εκτέλεσης μιας πρόβλεψης για ένα νέο πρότυπο [Specht (1990)].

Η απόδοση ενός ΠΝΔ επηρεάζεται σημαντικά από τις παραμέτρους λείανσης (smoothing parameters). Στην εργασία αυτή, προτείνεται ένα ΠΝΔ το οποίο ενσωματώνει διαφοροεξελικτικούς αλγόριθμους βελτιστοποίησης (Differential Evolution Algorithms, DE) [Storn (1997), Fan (2003)] για τη βελτιστοποίηση των παραμέτρων λείανσης του ΠΝΔ. Οι Διαφοροεξελικτικοί Αλγόριθμοι (ΔΕ) είναι μία οικογένεια στοχαστικών αλγορίθμων βελτιστοποίησης που βασίζονται σε πληθυσμούς ατόμων που αναπαριστούν πιθανές λύσεις στο χώρο αναζήτησης. Το προτεινόμενο σχήμα έχει εφαρμοστεί σε τρία σύνολα δεδομένων από το χώρο της βιοϊατρικής, το Wisconsin Breast Cancer Database, το Pima Indians Diabetes καθώς και το Heart από τη βάση δεδομένων Proben1 [Prechelt (1994)]. Επίσης τα αποτελέσματα αυτά έχουν συγκριθεί με τα αντίστοιχα αποτελέσματα που παρήχθησαν από νευρωνικά δίκτυα εμπρόσθιας τροφοδότησης.

## 2. ΠΙΘΑΝΟΤΙΚΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

Τα ΠΝΔ ανήκουν στην κατηγορία δικτύων με επίβλεψη (supervised) τα οποία χρησιμοποιούνται ευρέως στην αναγνώριση προτύπων καθώς και στην ταξινόμηση και ομαδοποίηση των δεδομένων [Specht (1994)]. Η τελική απόφαση ενός ΠΝΔ βασίζεται στον κανόνα ταξινόμησης του Bayes και τη μη παραμετρική εκτίμηση της συνάρτησης πυκνότητας πιθανότητας του Parzen [Parzen (1962), Specht (1990)]. Ένα σημαντικό προτέρημα των ΠΝΔ, σε σύγκριση με τα FNNs, είναι ότι προσφέρουν τρόπους ερμηνείας της δομής του δικτύου [Berthold *et al* (1998)], ενώ στα FNNs είναι δύσκολο να ερμηνευθεί τι αναπαριστά ο κάθε νευρώνας του δικτύου. Η τυπική διαδικασία εκπαίδευσης ενός

ΠΝΔ απαιτεί ένα μόνο πέρασμα όλων των προτύπων του συνόλου εκπαίδευσης [Specht (1990)]. Το χαρακτηριστικό αυτό μεταφράζεται σε πιο γρήγορη εκπαίδευση του ΠΝΔ από την εκπαίδευση ενός FNN.

Η δομή ενός ΠΝΔ είναι παρόμοια με αυτή ενός FNN μόνο που η αρχιτεκτονική του ΠΝΔ περιορίζεται σε τέσσερα επίπεδα, το *επίπεδο εισόδου, προτύπων, άθροισης και εξόδου*. Τα δεδομένα υπό μορφή διανυσμάτων  $X = (x_1, \dots, x_p)^T \in \mathbb{R}^p$  τοποθετούνται στους  $p$  νευρώνες εισόδου και στη συνέχεια περνούν στο επίπεδο προτύπων. Οι νευρώνες του επιπέδου προτύπων είναι χωρισμένοι σε  $K$  κλάσεις, όσες και οι κλάσεις στις οποίες είναι χωρισμένα τα δεδομένα. Η έξοδος του  $i$ -οστού νευρώνα προτύπων της  $k$ -οστής κλάσης υπολογίζεται με τη χρήση ενός κανονικού πυρήνα της μορφής:

$$(1) \quad f_{i,k}(X) = \frac{1}{(2\pi)^{p/2} \det(\Sigma)^{1/2}} \exp\left(-\frac{1}{2}(X - X_{i,k})^T \Sigma^{-1}(X - X_{i,k})\right),$$

όπου  $X_{i,k} \in \mathbb{R}^p$  είναι το κέντρο του πυρήνα και  $\Sigma$  ο  $p \times p$  πίνακας των παραμέτρων λείανσης. Ως  $X_{i,k}$  θεωρούμε το  $i$ -οστό διάνυσμα παρατηρήσεων της  $k$ -οστής κλάσης των δεδομένων. Θεωρούμε δυο εκδοχές του  $\Sigma$ :

(α) το ομοσκεδαστικό ΠΝΔ όπου  $\Sigma = \sigma^2 \cdot I_p$ , δηλαδή έχουμε μία ενιαία παράμετρο λείανσης

(β) το ετεροσκεδαστικό ΠΝΔ όπου  $\Sigma = \text{diag}\{\sigma_1^2, \sigma_2^2, \dots, \sigma_p^2\}$  δηλαδή ένας διαγώνιος πίνακας παραμέτρων.

Το επίπεδο άθροισης του δικτύου υπολογίζει τη δεσμευμένη πιθανότητα της κάθε κλάσης,

$$(2) \quad G_k(X) = \pi_k \cdot \sum_{i=1}^{M_k} f_{i,k}(X), \quad k \in \{1, \dots, K\},$$

όπου  $M_k$  είναι ο αριθμός των νευρώνων προτύπου της  $k$ -οστής κλάσης και  $\pi_k$  είναι η εκ των προτέρων πιθανότητά της. Η  $\pi_k$  ορίζεται από την αναλογία της κλάσης στο δείγμα ή αλλιώς θέτουμε  $\pi_k = 1/K$ . Ένα διάνυσμα  $X$  ταξινομείται σε εκείνη την κλάση στην οποία ο αντίστοιχος νευρώνας άθροισης αποκτά τη μεγαλύτερη τιμή, δηλαδή έχει τη μεγαλύτερη εκ των υστέρων πιθανότητα.

### 3. ΔΙΑΦΟΡΟΕΞΕΛΙΚΤΙΚΟΙ ΑΛΓΟΡΙΘΜΟΙ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ

Οι ΔΕ αλγόριθμοι είναι στοχαστικοί, αλγόριθμοι βελτιστοποίησης που βασίζονται σε πληθυσμούς [Storn (1997), Price (1999), Fan (2003)]. Οι αλγόριθμοι αυτοί εκκινούν με ένα πληθυσμό αρχικών πιθανών λύσεων και στη συνέχεια διερευνούν για «καλές» περιοχές λύσεων στο χώρο αναζήτησης. Οι πιθανές λύσεις καλούνται άτομα (individuals). Αν  $n$  είναι η διάσταση της βελτιστοποίησης, τότε ο ΔΕ αλγόριθμος χρησιμοποιεί  $n$ -διάστατα διανύσματα (πλήθους  $NP$ ) ως πληθυσμό της κάθε επανάληψης του αλγορίθμου, η οποία καλείται γενιά (generation). Σε κάθε γενιά επιδρούν δύο τελεστές σε κάθε άτομο, η μετάλλαξη (mutation) και ο επανασυνδυασμός (crossover), παράγοντας έτσι το νέο πληθυσμό. Αν  $g$  είναι η τρέχουσα γενιά, τότε για κάθε άτομο  $W_k^{(g)}$ ,  $k = 1, \dots, NP$ ,

το μεταλλαγμένο άτομο  $V_i^{(g+1)}$  καθορίζεται από μία από τις παρακάτω εξισώσεις (στρατηγικές):

- (1)  $V_i^{(g+1)} = W_{best}^{(g)} + F (W_{r_1}^{(g)} - W_{r_2}^{(g)})$ ,
- (2)  $V_i^{(g+1)} = W_{r_1}^{(g)} + F (W_{r_2}^{(g)} - W_{r_3}^{(g)})$ ,
- (3)  $V_i^{(g+1)} = W_i^{(g)} + F (W_{best}^{(g)} - W_i^{(g)}) + F (W_{r_1}^{(g)} - W_{r_2}^{(g)})$ ,
- (4)  $V_i^{(g+1)} = W_{best}^{(g)} + F (W_{r_1}^{(g)} - W_{r_2}^{(g)}) + F (w_{r_3}^{(g)} - W_{r_4}^{(g)})$ ,
- (5)  $V_i^{(g+1)} = W_{r_1}^{(g)} + F (W_{r_2}^{(g)} - W_{r_3}^{(g)}) + F (W_{r_4}^{(g)} - W_{r_5}^{(g)})$ ,
- (6)  $V_i^{(g+1)} = (W_{r_1}^{(g)} + W_{r_2}^{(g)} + W_{r_3}^{(g)}) / 3 + (p_2 - p_1) (W_{r_1}^{(g)} - W_{r_2}^{(g)}) + (p_3 - p_2) (W_{r_2}^{(g)} - W_{r_3}^{(g)}) + (p_1 - p_3) (W_{r_3}^{(g)} - W_{r_1}^{(g)})$ ,

όπου  $W_{best}^{(g)}$  είναι το καλύτερο άτομο της  $g$  γενιάς, δηλαδή αυτό που έχει τη μικρότερη τιμή στην αντικειμενική συνάρτηση,  $F \in [0, 2]$  είναι μια πραγματική σταθερά που καλείται σταθερά μετάλλαξης,  $CR \in [0, 1]$  είναι η σταθερά επανασυνδυασμού και  $r_1, r_2, r_3, r_4, r_5$  και  $r_6 \in \{1, 2, \dots, i-1, i+1, \dots, NP\}$  είναι τυχαίοι ακέραιοι που διαφέρουν ανά δύο. Οι παράμετροι  $F$  και  $CR$  θέτονται κάθε φορά σύμφωνα με τη στρατηγική που επιλέγεται. Τα  $p_1, p_2, p_3$ , υπολογίζονται ως εξής:

$$p_1 = |f(W_{r_1}^{(g)})|/p', \quad p_2 = |f(W_{r_2}^{(g)})|/p', \quad p_3 = |f(W_{r_3}^{(g)})|/p',$$

$$p' = |f(W_{r_1}^{(g)})| + |f(W_{r_2}^{(g)})| + |f(W_{r_3}^{(g)})|$$

Μετά τη φάση της μετάλλαξης, έπεται η φάση του επανασυνδυασμού. Για κάθε μεταλλαγμένο διάνυσμα  $V_i^{(g+1)}$ , παράγεται ένα τυχαίο διάνυσμα  $RV_i \in [0, 1]^n$  και γεννιέται το δοκιμαστικό διάνυσμα  $U_i^{(g+1)}$  σύμφωνα με:

$$u_{ij}^{(g+1)} = \begin{cases} v_{ij}^{(g+1)}, & \text{if } rv_{ij} \leq CR, \\ w_{ij}^{(g)}, & \text{διαφορετικά.} \end{cases}$$

Τέλος, το δοκιμαστικό άτομο  $U_i^{(g+1)}$  γίνεται μέλος του πληθυσμού αν προσφέρει κάποια μείωση στην συνάρτηση που ελαχιστοποιούμε, δηλαδή αν:

$$W_i^{(g+1)} = \begin{cases} U_i^{(g+1)}, & \text{if } f(U_i^{(g+1)}) < f(W_i^{(g)}), \\ W_i^{(g)}, & \text{διαφορετικά.} \end{cases}$$

#### 4. Η ΠΡΟΤΕΙΝΟΜΕΝΗ ΠΡΟΣΕΓΓΙΣΗ ΚΑΙ ΠΕΙΡΑΜΑΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ

Όπως αναφέρθηκε προηγουμένως ο πίνακας παραμέτρων λείανσης  $\Sigma$  παίζει σημαντικό ρόλο στην αποδοτικότητα του ΠΠΔ. Στην εργασία αυτή, ο καθορισμός του πίνακα παραμέτρων  $\Sigma$  επιτυγχάνεται μέσω των διαφόρων στρατηγικών των διαφοροεξελικτικών αλγορίθμων. Οι ΔΕ αλγόριθμοι αρχικοποιούνται τυχαία με ένα πληθυσμό τιμών του

ΠΙΝΑΚΑΣ 1: Ποσοστό επιτυχούς ταξινόμησης του ΔΕ-ΠΙΝΔ για το WBCD

Μοντέλο	Στρατηγική	Μέση Επιτ. %	Τυπ. Απ.	Min.	Max.
Ομοσκεδ. ΔΕ-ΠΙΝΔ	1 <sup>η</sup>	96,32	0,43	95,98	97,70
Ετεροσκ. ΔΕ-ΠΙΝΔ	3 <sup>η</sup>	97,67	0,60	97,13	98,85
FNN		98,53	0,60		

Σ και η βελτιστοποίηση επιτυγχάνεται για ένα μέρος του συνόλου δεδομένων ως προς το ποσοστό επιτυχούς ταξινόμησης. Το μοντέλο αυτό το ονομάζουμε *Διαφοροεξελικτικό Πιθανοτικό Νευρωνικό Δίκτυο* (ΔΕ-ΠΙΝΔ).

Εφαρμόσαμε το ΔΕ-ΠΙΝΔ σε τρία γνωστά σύνολα δεδομένων από το χώρο της βιοϊατρικής, συγκεκριμένα το Wisconsin Breast Cancer Database, (WBCD), το Pima Indians Diabetes, (PID) και Heart. Τα δεδομένα που χρησιμοποιήθηκαν, προέρχονται από τη βάση δεδομένων Proben1 και έχουμε ακολουθήσει τις προτεινόμενες οδηγίες. Για να καθοριστεί ο πίνακας Σ, ένας πληθυσμός από 10 άτομα (δηλαδή 10 πίνακες Σ) εξελίχθηκε για 50 γενιές για την περίπτωση του ομοσκεδαστικού ΔΕ-ΠΙΝΔ, και πληθυσμός 20 ατόμων εξελίχθηκε για 100 γενιές για το ετεροσκεδαστικό. Επίσης η αποδοτικότητα του ΔΕ-ΠΙΝΔ συγκρίθηκε με αυτή των FNNs. Όλα τα αποτελέσματα των FNNs που παραθέτονται είναι τα καλύτερα αποτελέσματα που έχουν επιτευχθεί από το Proben1.

Παρακάτω παρουσιάζεται μια σύντομη περιγραφή των δεδομένων. Το πρώτο σύνολο δεδομένων είναι το WBCD και ο στόχος είναι να προβλεφθεί αν ένας όγκος στο μαστό είναι καλοήγησ ή όχι. Υπάρχουν 9 μεταβλητές (κυτταρικές μετρήσεις) και το δείγμα αποτελείται από 699 ασθενείς. Το δεύτερο σύνολο δεδομένων είναι το PID, στο οποίο ο στόχος είναι να διαγνωσθεί η ασθένεια του διαβήτη στους ινδιάνους Pima. Το μέγεθος του δείγματος είναι 768 άτομα και οι μεταβλητές είναι 8. Τέλος, ο στόχος του τρίτου συνόλου δεδομένων (Heart) είναι η πρόβλεψη ασθενειών καρδιάς, όπου έχουμε 920 άτομα και 35 μεταβλητές (στην υλοποίηση χρησιμοποιήθηκαν μόνο οι 32 μεταβλητές, διότι οι υπόλοιπες ήταν μηδενικές). Και στα τρία προβλήματα, η εξαρτημένη μεταβλητή είναι δίτημη.

ΠΙΝΑΚΑΣ 2: Ποσοστό επιτυχούς ταξινόμησης του ΔΕ-ΠΙΝΔ για το PID

Μοντέλο	Στρατηγική	Μέση Επιτ. %	Τυπ. Απ.	Min.	Max.
Ομοσκεδ. ΔΕ-ΠΙΝΔ	2 <sup>η</sup> , 3 <sup>η</sup> , 6 <sup>η</sup>	71,35	0,00	71,35	71,35
Ετεροσκ. ΔΕ-ΠΙΝΔ	1 <sup>η</sup>	75,76	0,92	74,48	78,12
FNN		75,43	3,53		

Κάθε μοντέλο εκτελέστηκε για διάφορες αρχικές τιμές 20 φορές για κάθε ένα πρόβλημα. Στους Πίνακες 1 – 3 παρουσιάζεται ο μέσος όρος, η τυπική απόκλιση καθώς και η ελάχιστη και μέγιστη τιμή των ποσοστών της επιτυχούς ταξινόμησης (σε μορφή ποσοστού) για τα τρία σύνολα δεδομένων. Για κάθε περίπτωση παρουσιάζεται η στρατηγική των διαφοροεξελικτικών αλγορίθμων που απέδωσε το καλύτερο αποτέλεσμα, καθώς και

ΠΙΝΑΚΑΣ 3: Ποσοστό επιτυχούς ταξινόμησης του ΔΕ-ΠΝΔ για το Heart

Μοντέλο	Στρατηγική	Μέση Επιτ.%	Τυπ.Απ.	Min.	Max.
Ομοσκεδ. ΔΕ-ΠΝΔ	2 <sup>η</sup> - 6 <sup>η</sup>	80,43	0,00	80,43	80,43
Ετεροσκ. ΔΕ-ΠΝΔ	6 <sup>η</sup>	81,15	1,00	79,57	83,04
FNN		80,11	2,27		

τα αντίστοιχα αποτελέσματα των FNNs. Μπορούμε να παρατηρήσουμε ότι το ετεροσκεδαστικό ΔΕ-ΠΝΔ απέδωσε καλύτερα σε σχέση με το ομοσκεδαστικό ΔΕ-ΠΝΔ. Επίσης, τα FNNs απέδωσαν καλύτερα μόνο στο WBCD σε σχέση με τα ΔΕ-ΠΝΔ. Ακόμα, και στα τρία προβλήματα η τυπική απόκλιση του ποσοστού επιτυχούς ταξινόμησης ήταν μικρότερη στα ΔΕ-ΠΝΔ σε σχέση με τα FNNs. Ιδιαίτερα στα ομοσκεδαστικά ΔΕ-ΠΝΔ στα οποία είχαμε μηδενική διασπορά, ο αλγόριθμος συγκλίνει σε όλες τις εκτελέσεις σχεδόν στο ίδιο σημείο, γεγονός που δηλώνει ότι πρόκειται για ένα αξιόπιστο σχήμα.

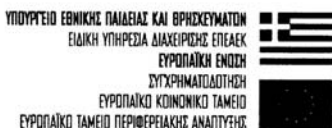
Εφαρμόζοντας έναν t-έλεγχο για να συγκρίνουμε την «καλύτερη» μέση απόδοση των ΔΕ-ΠΝΔ και FNNs στο WBCD, συμπεραίνουμε ότι υπάρχει στατιστικά σημαντική διαφορά υπέρ των FNNs ( $p$ -value < 0.001). Για τα δύο άλλα προβλήματα, χρησιμοποιήθηκε ο t-έλεγχος για άνισες διασπορές, καθώς υπάρχει μεγάλη διαφορά μεταξύ των διασπορών των FNNs και ΔΕ-ΠΝΔ. Δεν παρατηρήθηκε στατιστικά σημαντική διαφορά μεταξύ των ΔΕ-ΠΝΔ και FNNs σε κανένα από τα άλλα δύο προβλήματα ( $p$ -value = 0.682 για το PID και  $p$ -value = 0.051 για το Heart). Συγκρίνοντας την καλύτερη απόδοση του ομοσκεδαστικού με αυτή του ετεροσκεδαστικού ΔΕ-ΠΝΔ στο WBCD συμπεραίνουμε ότι υπάρχει στατιστικά σημαντική διαφορά υπέρ του ετεροσκεδαστικού ( $p$ -value < 0.001). Και στα άλλα δύο προβλήματα, υπερείχε το ετεροσκεδαστικό ΔΕ-ΠΝΔ αλλά δεν έγιναν στατιστικοί έλεγχοι καθώς η διασπορά είναι μηδενική.

## 5. ΣΥΜΠΕΡΑΣΜΑΤΑ

Στην παρούσα εργασία προτάθηκε ένα αυτοπροσαρμοζόμενο σχήμα για Πιθανοτικά Νευρωνικά Δίκτυα. Η προτεινόμενη προσέγγιση ενσωματώνει τους Διαφοροεξελικτικούς αλγόριθμους βελτιστοποίησης για αναζήτηση ενός κατάλληλου πίνακα παραμέτρων λείανσης του ΠΝΔ έτσι ώστε να μεγιστοποιηθεί το ποσοστό επιτυχούς ταξινόμησης. Η αποτελεσματικότητα του προτεινόμενου μοντέλου δοκιμάστηκε σε τρία σύνολα δεδομένων από το χώρο της βιοϊατρικής όπως η «βάση δεδομένων του Wisconsin για τον καρκίνο του μαστού», η «ασθένεια διαβήτη για του ινδιάνους Pima» καθώς και η «πρόγνωση ασθενειών καρδιάς» με ενθαρρυντικά αποτελέσματα. Χρησιμοποιήθηκαν δύο εκδοχές των ΠΝΔ, η πρώτη με μία μόνο παράμετρο λείανσης στους πυρήνες του ΠΝΔ και η δεύτερη με διαγώνιο πίνακα παραμέτρων λείανσης. Το ετεροσκεδαστικό ΔΕ-ΠΝΔ έδωσε καλύτερα αποτελέσματα σε σχέση με το ομοσκεδαστικό και στα τρία προβλήματα. Για την περαιτέρω διερεύνηση των δυνατοτήτων του, το ΔΕ-ΠΝΔ συγκρίθηκε με τεχνητά νευρωνικά δίκτυα εμπρόσθιας τροφοδότησης με ικανοποιητικά αποτελέσματα. Μελλοντικά, θα εξεταστεί και η περίπτωση όπου ο πίνακας  $\Sigma$  δεν θα είναι διαγώνιος, καθώς η βελτιστοποίηση γίνεται περισσότερο περίπλοκη και πρέπει να ληφθεί υπόψη η υπερπροσαρμογή (overfitting) του μοντέλου στα δεδομένα.

## 6. ΕΥΧΑΡΙΣΤΙΕΣ

Ευχαριστούμε το Επιχειρησιακό Πρόγραμμα Εκπαίδευσης και Αρχικής Επαγγελματικής Κατάρτισης II (ΕΠΕΑΕΚ II) και συγκεκριμένα το πρόγραμμα ΗΡΑΚΛΕΙΤΟΣ για τη χρηματοδότηση της παραπάνω εργασίας.



## ABSTRACT

In this paper a new model called Differential Evolution Probabilistic Neural Network (DE-PNN) is proposed. The simple Probabilistic Neural Network (PNN) is a model employed for classification and pattern recognition tasks into predefined classes and is also known in the statistical literature as Kernel Discriminant Analysis. This model incorporates the Bayes decision rule, which means that it classifies the new observations according to their class posterior probabilities.

The proposed DE-PNNs incorporate differential evolution algorithms in order to optimize the smoothing parameters of the kernels of the PNN. Differential evolution optimization algorithms are population-based algorithms that do not require computations of derivatives but only functional computations. The proposed DE-PNNs are applied to three biomedical real-world applications and are compared to FNNs with encouraging results.

## ΑΝΑΦΟΡΕΣ

- Berthold, M. and J. Diamond (1998). Constructive Training of Probabilistic Neural Networks. *Neurocomputing*, 167–183.
- Fan, H.Y. and J. Lampinen (2003). A Trigonometric Mutation Operation to Differential Evolution. *Journal of Global Optimization*, 27, 105–129.
- Georgiou V.L., N.G. Pavlidis, K.E. Parsopoulos, Ph.D. Alevizos and M. N. Vrahatis (2006). New Self-Adaptive Probabilistic Neural Networks in Bioinformatic and Medical Tasks. *Int. Journal on Artificial Intelligence Tools* (to appear).

- Georgiou V.L., N.G. Pavlidis, K.E. Parsopoulos, Ph.D. Alevizos and M. N. Vrahatis (2004). Optimizing the Performance of Probabilistic Neural Networks in a Bioinformatics Task. In *Proceedings of the EUNITE 2004 Conference*, 34–40.
- Georgiou V.L., N.G. Pavlidis, K.E. Parsopoulos, Ph.D. Alevizos and M. N. Vrahatis (2004). Evolutionary Adaptive Schemes of Probabilistic Neural Networks. In *Lecture Series on Computer and Computational Sciences, ICCMSE 2004*, 1, 852–855.
- Huang, C.J. (2002). A performance analysis of cancer classification using feature extraction and probabilistic neural networks. In *Proceedings of the 7th Conference on Artificial Intelligence and Applications*, 374–378.
- Orr, R.K. (1997). Use of a probabilistic neural network to estimate the risk of mortality after cardiac surgery. *Medical Decision Making*, 17(2), 178–185.
- Parzen, E. (1962). On the estimation of a probability density function and mode. *Annals of Mathematical Statistics*, 3, 1065–1076.
- Prechelt, L. (1994). Proben1: A set of neural network benchmark problems and benchmarking rules. Technical Report 21/94, Fakultät für Informatik, Universität Karlsruhe. ([/afs/cs/project/connect/bench/contrib/prechelt/proben1.tar.gz](#))
- Price, K. (1999). An introduction to differential evolution. *New Ideas in Optimization*, McGraw-Hill, London (UK), 79–108.
- Specht, D.F. (1990). Probabilistic Neural Networks. *Neural Networks*, 1(3), 109–118.
- Specht, D.F. and H. Romsdahl (1994). Experience with Adaptive Probabilistic Neural Network and Adaptive General Regression Neural Network. *Proceedings of the IEEE International Conference on Neural Networks*, 2, 1203–1208.
- Storn, R. and K. Price (1997). Differential Evolution—A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces. *Journal of Global Optimization*, 11, 341–359.
- Tan, A.C. and D. Gilbert (2003). An Empirical Comparison of Supervised Machine Learning Techniques in Bioinformatics, *Proceedings of the 1st Asia Pacific Bioinformatics Conference (APBC 2003)*, 219–222.
- Y. Wang, T. Adali, S. Kung, and Z. Szabo (1998). Quantification and segmentation of brain tissues from mr images: A probabilistic neural network approach. *IEEE Transactions on Image Processing*, 7(8), 1165–1181.