

Εφαρμογή του Νευρωνικού Δικτύου Πολυεπίπεδο Perceptron στην Κατηγοριοποίηση Ιατρικών Δεδομένων

Θεοφάνης Γεροδήμος
ma11553@uoi.gr
Τμήμα Μαθηματικού,
Πανεπιστήμιο Ιωαννίνων

Περίληψη-Καθημερινά, παράγεται ένας τεράστιος όγκος ιατρικών δεδομένων. Η μηχανική μάθηση είναι μια πολύ αποτελεσματική τεχνική που χρησιμοποιείται στην ανάλυση των δεδομένων και στην εύρεση προτύπων στον τομέα της λήψης αποφάσεων που σχετίζονται με την Ιατρική Επιστήμη. Επειδή οι αποφάσεις στον ιατρικό τομέα ασχολούνται με την έκβαση της υγείας των ασθενών, απαιτείται ένα υψηλό επίπεδο ακριβείας στις προβλέψεις από τους αλγόριθμους μηχανικής μάθησης που εφαρμόζονται. Η εργασία αυτή παρουσιάζει το πολυεπίπεδο perceptron (MLPNN) και πως αυτό εφαρμόζεται στην Ιατρική Επιστήμη, πιο συγκεκριμένα στον τομέα της καρδιολογίας.

Λέξεις κλειδιά: Υπολογιστική Νοημοσύνη, Πολυεπίπεδο Perceptron,

Τεχνητά Νευρωνικά Δίκτυα,

1 Εισαγωγή

Οι τεχνικές εξόρυξης δεδομένων και μηχανικής μάθησης μπορούν να χρησιμοποιηθούν στον τομέα της υγειονομικής περίθαλψης, ειδικά σε σύνολα δεδομένων που σχετίζονται με καρδιαγγειακές παθήσεις. Διαφορετικές τεχνικές εξόρυξης δεδομένων έχουν χρησιμοποιηθεί στο παρελθόν για την εξαγωγή πληροφοριών και προτύπων που συμβάλλουν στη βελτίωση της ιατρικής λήψης αποφάσεων [6]. Οι προσεγγίσεις μηχανικής μάθησης μπορούν να ωφελήσουν την υγειονομική περίθαλψη μέσω διαφόρων προσεγγίσεων, όπως μείωση της διάρκειας θεραπείας, ανίχνευση αιτιών και συμπτωμάτων της νόσου κλπ [7].

2 Νευρωνικά Δίκτυα

Η τεχνική των νευρωνικών δικτύων είναι ένα σημαντικό εργαλείο για την ταξινόμηση προτύπων που μπορεί να αυτοπροσαρμοστεί στα δεδομένα, ανεξάρτητα από οποιαδήποτε ειδική λειτουργική απαίτηση του πρωταρχικού μοντέλου [4].

Σε παλαιότερη εργασία, έγινε χρήση νευρωνικού δικτύου πολλαπλών επιπέδων perceptron (MLPNN) με τον αλγόριθμο οπισθοδρόμησης (BP) για την πρόβλεψη καρδιακών παθήσεων. Οι συγγραφείς απέδειξαν ότι τα νευρωνικά δίκτυα έχουν σημαντικές δυνατότητες στην ακρίβεια της διάγνωσης [2]. Η οπισθοδρόμηση μπορεί να χρησιμοποιηθεί μαζί με μια μέθοδο βελτιστοποίησης όπως η κάθοδος βασισμένη στην κλίση (gradient descent) για τη φάση εκπαίδευσης των νευρωνικών δικτύων. Με αυτήν τη μέθοδο υπολογίζεται η κλίση μιας συνάρτησης απώλειας σε σχέση με όλα τα βάρη στο δίκτυο [5]. Ωστόσο, το πολυεπίπεδο perceptron (MLP) είναι ένα νευρωνικό δίκτυο τροφοδοσίας που χαρτογραφεί τις εισόδους σε ένα κατάλληλο σύνολο εξόδων. Το MLP αποτελείται από πολλά επίπεδα κόμβων μέσα σε ένα γράφημα, έτσι ώστε κάθε επίπεδο να συνδέεται πλήρως με το επόμενο με μια μη γραμμική συνάρτηση ενεργοποίησης, εξαιρουμένων των κόμβων εισόδου. Το MLP χρησιμοποιεί μια τεχνική εκμάθησης με επίβλεψη που ονομάζεται εκπαίδευση οπισθοδρομικής διάδοσης σφάλματος και μια μη γραμ-

μική συνάρτηση ενεργοποίησης [3]. Σε μια άλλη μελέτη, οι συγγραφείς, χρησιμοποιώντας νευρωνικά δίκτυα σε ένα σύνολο δεδομένων που αποτελείται από 573 εγγραφές, χωρισμένα στα δύο μέρη, της εκπαίδευσης και του ελέγχου, ισχυρίστηκαν ότι είχαν 100% ακρίβεια αυξάνοντας τον αριθμό των χαρακτηριστικών από 13 σε 15 [1].

3 (MLP) Πολυεπίπεδο Perceptron

Τα MLPs έχουν εφαρμοσθεί σε πολλά δύσκολα προβλήματα και εκπαιδεύονται με επίβλεψη (supervised learning), με τον γνωστό αλγόριθμο οπισθοδρομικής διάδοσης του σφάλματος (error back propagation). Η διαδικασία εκπαίδευσης στον αλγόριθμο back-propagation περιλαμβάνει υπολογισμούς που υλοποιούνται σε δύο περάσματα. Ένα πέρασμα κατά την ευθεία φορά και ένα πέρασμα κατά την αντίστροφη φορά. Στο ευθύ πέρασμα εφαρμόζεται ένα πρότυπο στις εισόδους του δικτύου, πραγματοποιούνται οι υπολογισμοί κατά την ορθή φορά και τέλος παράγεται ένα σύνολο από εξόδους που αποτελούν την πραγματική έξοδο του δικτύου. Αντίθετα κατά τη διάρκεια του αντίστροφου περάσματος, τα βάρη προσαρμόζονται σύμφωνα με τον κανόνα διόρθωσης σφάλματος (error correction). Συγκεκριμένα η πραγματική τιμή εξόδου αφαιρείται από την αντίστοιχη επιθυμητή και παράγεται το σήμα σφάλματος.

Κάθε κρυμμένος νευρώνας περ-

ιέχει μια μη γραμμική συνάρτηση ενεργοποίησης (activation function)

όπου $u_j(n)$ καθορίζεται από την σχέση:

$$y = \frac{1}{1 + e^{-u}}$$

$$u_j(n) = \sum_{i=1}^p w_{ji}(n)y_i(n)$$

3.1 Τα δύο περάσματα του αλγόριθμου

Στην εφαρμογή του αλγόριθμου μπορούμε να διακρίνουμε δύο περάσματα, το πέρασμα ορθής φοράς (forward pass), και το αντίστροφης φοράς (reverse pass). Στο ευθύ πέρασμα το σήμα που εμφανίζεται στην έξοδο του j -νευρώνα δίνεται από την σχέση:

$$y_j(n) = \phi(u_j(n))$$

όπου p : ο αριθμός εισόδων που εφαρμόζονται στο νευρώνα j και $y_j(n)$: το σήμα εισόδου του νευρώνα j . Διαφορίζοντας την λογιστική συνάρτηση έχουμε:

$$\begin{aligned} y' &= \frac{\theta y}{\theta u} = \phi'_j(u) = \frac{e^{-u}}{(1 + e^{-u})^2} \\ &= y(1 - y) \end{aligned}$$

Επομένως ορίζουμε την τοπική κλίση για τον j νευρώνα ως:

$$\delta_j(n) =$$

References

- [1] Chaitrali Dangare and Sulabha Apte. A data mining approach for prediction of heart disease using neural networks. *International Journal of Computer Engineering and Technology (IJCET)*, 3(3), 2012.
- [2] Curt G DeGroff, Sanjay Bhatikar, Jean Hertzberg, Robin Shandas, Lillian Valdes-Cruz, and Roop L Mahajan. Artificial neural network-based method of screening heart murmurs in children. *Circulation*, 103(22):2711–2716, 2001.
- [3] Don R Hush and Bill G Horne. Progress in supervised neural networks. *IEEE signal processing magazine*, 10(1):8–39, 1993.
- [4] Ling Liu. From data privacy to location privacy: models and algorithms. In *Proceedings of the 33rd international conference on Very large data bases*, pages 1429–1430. Citeseer, 2007.
- [5] Alexander G Ororbia II, C Lee Giles, and Daniel Kifer. Unifying adversarial training algorithms with flexible deep data gradient regularization. *arXiv preprint arXiv:1601.07213*, 2016.

- [6] Umair Shafique, Fiaz Majeed, Haseeb Qaiser, and Irfan Ul Mustafa. Data mining in healthcare for heart diseases. *International Journal of Innovation and Applied Studies*, 10(4):1312, 2015.
- [7] Jyoti Soni, Ujma Ansari, Dipesh Sharma, and Sunita Soni. Predictive data mining for medical diagnosis: An overview of heart disease prediction. *International Journal of Computer Applications*, 17(8):43–48, 2011.