

Εκπαίδευση Νευρωνικού Δικτύου με Backpropagation

Εργασία 5

Μπεκιάρης Θεοφάνης ΑΕΜ:8200

Στόχος

Ο στόχος της εργασίας είναι η δημιουργία ενός νευρωνικού δικτύου και η εκπαίδευση του με την χρήση του αλγόριθμου backpropagation όπως παρουσιάζεται στις διαφάνειες του μαθήματος στην διάλεξη 11.

Περιεχόμενα

Η λογική των προγραμμάτων που περιέχονται στην εργασία στηρίζονται στην ίδια λογική με τα παραδείγματα του perceptron από τις διαλέξεις, δηλαδή δημιουργία μίας συνάρτησης classification για τον καθορισμό των δεδομένων διάταξης και την δημιουργία των αντίστοιχων γραφημάτων επίδειξης, μία συνάρτηση train που χρησιμοποιεί τα μισά δεδομένα για την εκπαίδευση του δικτύου και μια συνάρτηση classify που χρησιμοποιεί τα υπόλοιπα δεδομένα για να τα κατηγοριοποιήσει με την χρήση του προηγούμενου δικτύου που εκπαιδεύσαμε ώστε τελικά να επιβεβαιώσουμε την επιτυχία της διαδικασίας.

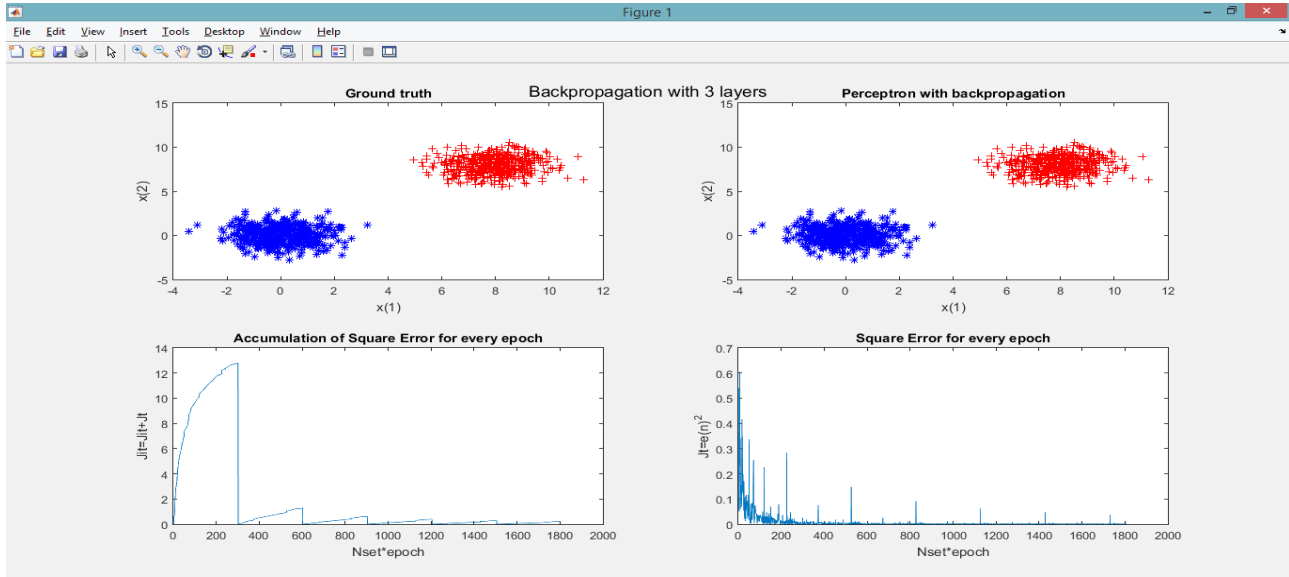
Φάκελος 3LayersSingleData: Στον φάκελο περιέχονται οι παραπάνω συναρτήσεις που υλοποιούν την κατηγοριοποίηση δεδομένων δομής όπως αυτής του παραδείγματος perceptron από τις διαλέξεις, δηλαδή δισδιάστατων διανυσμάτων εισόδου με απλή ετικέτα μίας διάστασης και δύο κατηγορίες δεδομένων. Το δίκτυο αποτελείται από 3 στρώματα, το πρώτο είναι της εισόδου, το δεύτερο ένα κρυφό στρώμα και το τρίτο και τελευταίο της εξόδου.

Φάκελος 3LayersVectors: Οι αντίστοιχες συναρτήσεις του φακέλου υλοποιούν ένα δίκτυο 3 στρωμάτων όπως και πριν, τώρα όμως τα δεδομένα κατηγοριοποίησης είναι αυτά που προκύπτουν από την συνάρτηση $[x \ y] = \text{iris_dataset}$ του MATLAB, όπου το x περιέχει τα διανύσματα εισόδου τεσσάρων διαστάσεων και το y τις τρισδιάστατες ετικέτες των 3 κατηγοριών που πρέπει να κατηγοριοποιηθούν τα δεδομένα.

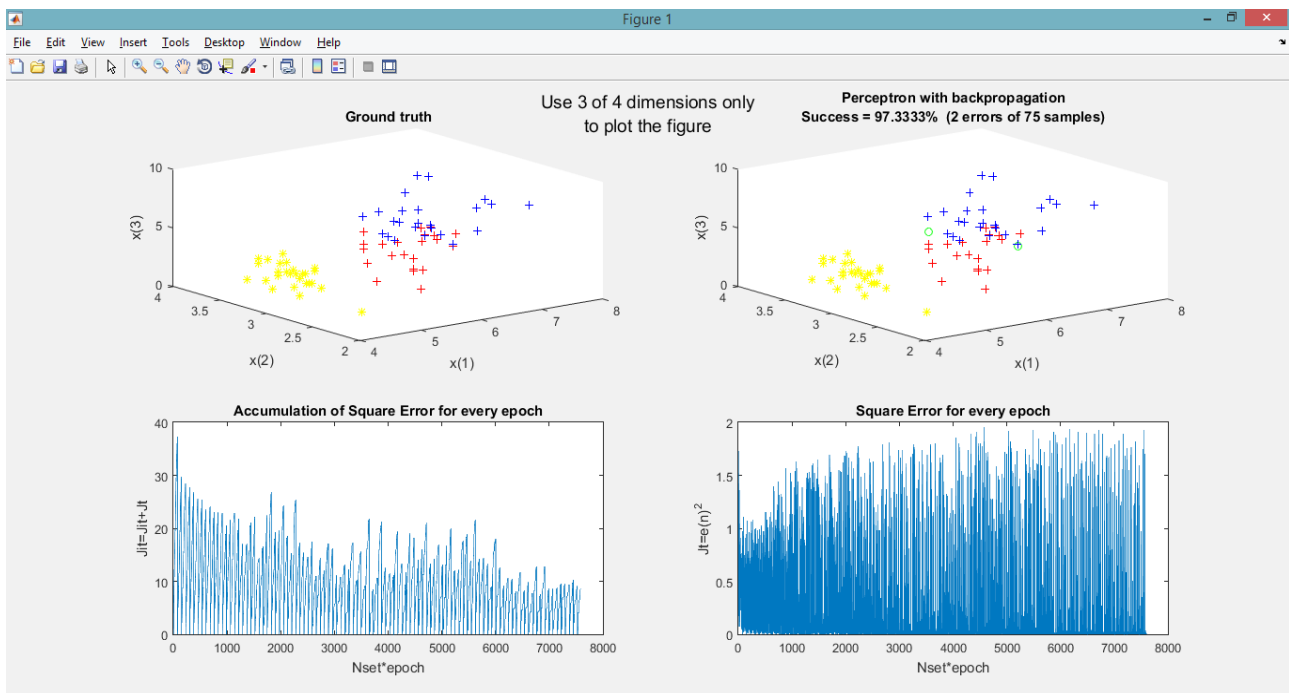
Φάκελος MultilayersVectors: Οι αντίστοιχες συναρτήσεις του φακέλου κατηγοριοποιούν τα δεδομένα από την συνάρτηση $[x \ y] = \text{iris_dataset}$ του MATLAB όμως αυτή τη φορά ο αριθμός των κρυφών στρωμάτων του δικτύου ορίζεται από τον χρήστη μέσω της μεταβλητής hidden_layers μέσα στην συνάρτηση backpropagation_train. Οι πίνακες βαρών των κρυφών στρωμάτων είναι τετραγωνικοί nxn για να ταιριάζουν στα αντίστοιχα στρώματα εισόδου και εξόδου. Η προκαθορισμένη τιμή της μεταβλητής είναι 3, δηλαδή μαζί με τα στρώματα εισόδου και εξόδου έχουμε ένα δίκτυο 5 στρωμάτων.

Γραφήματα και αποτελέσματα

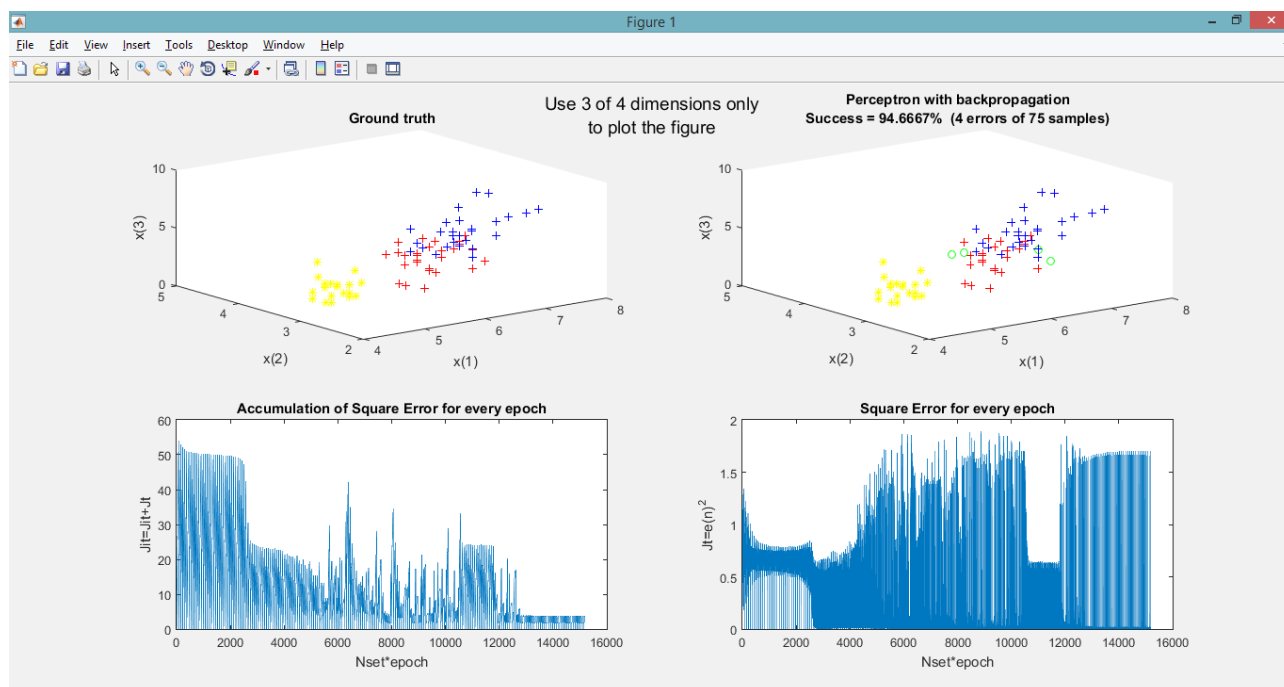
A) Εκμάθηση δικτύου 3 στρωμάτων με τον backpropagation για 6 εποχές, αλγόριθμος φακέλου 3LayersSingleData. Το πάνω αριστερό διάγραμμα είναι η πραγματική κατηγοριοποίηση των δεδομένων, το πάνω δεξιά η κατηγοριοποίηση από το δίκτυο, κάτω αριστερά η συσσώρευση του τετραγωνικού σφάλματος εξόδου του δικτύου (για κάθε εποχή) και κάτω δεξιά το τετραγωνικό σφάλμα (για κάθε εποχή).



B) Στο παρακάτω διάγραμμα φαίνονται τα αποτελέσματα του αλγορίθμου του φακέλου 3LayersVectors για την κατηγοριοποίηση των διανυσμάτων της συνάρτησης $[x \ y] = \text{iris_dataset}$. Για να μπορέσουμε να δημιουργήσουμε ένα εποπτικό διάγραμμα με στόχο την παρατήρηση των αποτελεσμάτων από τον αλγόριθμο, έχουν χρησιμοποιηθεί οι 3 πρώτες τιμές από κάθε διάνυσμα των τεσσάρων διαστάσεων ώστε να αντιπροσωπεύουν ένα σημείο του χώρου με στόχο την δημιουργία ενός διαγράμματος όπως το παραπάνω. Να τονιστεί προς αποφυγή οποιασδήποτε παρεξήγησης ότι ο αλγόριθμος λαμβάνει υπόψη όλες τις διαστάσεις των διανυσμάτων εισόδου (και τις 4), απλώς για την γραφική απεικόνιση χρησιμοποιούμε τις τρεις πρώτες διαστάσεις όπου το κάθε σημείο αντιπροσωπεύει το αντίστοιχο 4D διάνυσμα και ανάλογα το σύμβολο με το οποίο χρωματίζεται δείχνει την επιτυχία του υπολογισμού του αντίστοιχου διανύσματος.



Γ)Ομοίως με παραπάνω φαίνεται και το διάγραμμα για τον τρίτο αλγόριθμο του φακέλου MultilayersVectors. Εδώ έχουμε συνολικά 5 στρώματα.



Παρατηρήσεις

Στο πρώτο γράφημα βλέπουμε τα αποτελέσματα από τον αλγόριθμο για δίκτυο 2 διαστάσεων εισόδου και 1 διάστασης εξόδου. Παρατηρούμε ότι η επιτυχία του αλγόριθμου είναι αρκετά μεγάλη για τα δεδομένα. Στα διαγράμματα συσσώρευσης τετραγωνικού σφάλματος και τετραγωνικού σφάλματος φαίνεται η μείωση της συνάρτησης της κάθε εποχής. Η συσσώρευση μειώνεται για κάθε εποχή με εμφανή τρόπο καθώς και το σφάλμα. Στο δεύτερο και τρίτο διάγραμμα τα αποτελέσματα του αλγορίθμου φαίνονται καλά από τα 2 πάνω διαγράμματα καθώς και από το ποσοστό επιτυχίας που αναγράφεται (97.3% και 94.6% αντίστοιχα). Το διάγραμμα σφάλματος δεν είναι τόσο ευκρινές αλλά από το διάγραμμα συσσώρευσης μπορούμε να καταλάβουμε την μείωση του τετραγωνικού σφάλματος που λαμβάνει χώρα συνάρτησης της κάθε εποχής (100 εποχές για το πρώτο 200 για το δεύτερο). Αξίζει να αναφέρω ότι μετά από δοκιμές κατάλαβα ότι ο αριθμός των εσωτερικών στρωμάτων όσο μεγαλώνει απαιτεί περισσότερες εποχές για ικανοποιητικά αποτελέσματα και όσο μεγαλώνει ο αριθμός των στρωμάτων δεν σημαίνει ότι αλγόριθμος γίνεται καλύτερος. Χρειάζονται δοκιμές και παρατήρηση για να εξαχθούν οι κατάλληλες αναλογίες. Το ίδιο ισχύει και για τον αριθμό των νευρώνων, θα έλεγα ότι για πολυδιάστατα δεδομένα εισόδου εξόδου με μεγάλο αριθμό εισόδων και εξόδων ο αριθμός των νευρώνων δεν θα πρέπει να είναι πολύ μικρός σε σχέση με τον αριθμό των εισόδων-εξόδων καθώς είναι σχεδόν αδύνατο να καθοριστούν οι εξοδοί αν αυτές είναι πολλές. Επίσης για μεγάλο αριθμό νευρώνων γίνεται δύσκολη η εκμάθηση και χρειάζονται μεγαλύτερος αριθμός εποχών.