

Υπολογιστική Νοημοσύνη

Εργαστηριακές Ασκήσεις 2021

Αναφορά

Τζώρτζης Ευάγγελος

3088

Άσκηση 1:

Σε αυτή την άσκηση υλοποιήθηκαν δύο τύποι νευρωνικών δικτύων με 2 και 3 επίπεδα κρυμμένων νευρώνων.

Έγιναν πειράματα με διάφορους συνδυασμούς αριθμού νευρώνων και συναρτήσεων ενεργοποίησης.

Η επιλογή του ρυθμού μάθησης σε κάθε πείραμα έγινε κατάλληλα ώστε να υπάρχει μείωση του λάθους και όχι ταλάντωση στις τιμές του (παρατηρήθηκαν ταλαντώσεις αντί για μείωση όταν ο ρυθμός μάθησης ήταν μεγαλύτερος από κάποια τιμή).

Ακόμα το κατώφλι λάθους επιλέχθηκε κατάλληλα για τον τερματισμό των προγραμμάτων σε ένα συγκεκριμένο χρονικό φάσμα 30 δ. έως 4 λ. με κάποιες διακυμάνσεις.

Παρακάτω υπάρχει ένας πίνακας με κάποια πειράματα που έγιναν.

Η λογιστική συνάρτηση στο επίπεδο εξόδου είχε καλύτερα αποτελέσματα σε σχέση με την \tanh .

Όταν χρησιμοποιήθηκε η relu σε κάποιο κρυμμένο επίπεδο τα ποσοστά επιτυχίας ήταν κατώτερα από τις αντίστοιχες περιπτώσεις όπου όλα τα επίπεδα είχαν ως συνάρτηση ενεργοποίησης την \tanh .

Στα πειράματα στα οποία το $B = N/10$ τα αποτελέσματα των δοκιμών με το test set ήταν παρόμοια, έχοντας ένα μικρό προβάδισμα, με αυτά όταν το $B = N/100$. Η διαφορά μεταξύ τους είναι ότι για παρόμοιο ποσοστό επιτυχίας, ο χρόνος εκπαίδευσης ήταν μεγαλύτερος για $B = N/10$.

Παρατηρήθηκε ως βέλτιστο το πρόγραμμα Π1 με 9 νευρώνες και συνάρτηση ενεργοποίησης την \tanh και στα 2 κρυμμένα επίπεδα και τη λογιστική συνάρτηση στο επίπεδο εξόδου, με batch size = $N/10$ σε 230 δευτερόλεπτα με ποσοστό σωστής πρόβλεψης στα δοκιμαστικά δεδομένα 97.375%. Αρκετά κοντά υπήρξαν και εκτελέσεις του Π2 με διάφορες να έχουν ποσοστά ~95-96%.

Από τα διάφορα πειράματα βγαίνει το συμπέρασμα από το 3^ο κρυμμένο επίπεδο δεν υπάρχει βελτίωση στη γενικευτική ικανότητα των δικτύων και οι χρόνοι εκπαίδευσης είναι λίγο μεγαλύτεροι στα δίκτυα με 3 κρυμμένα επίπεδα, όμως υπάρχουν περισσότερα πειράματα με καλά αποτελέσματα (>90% σωστές προβλέψεις) στις διάφορες περιπτώσεις αριθμού νευρώνων και συναρτήσεων ενεργοποίησης. Τέλος τα δίκτυα με 3 κρυμμένα επίπεδα έχουν καλύτερα αποτελέσματα και για μικρότερο αριθμό νευρώνων.

Άρα το όφελος από τη χρήση του 3^{ου} κρυμμένου επιπέδου είναι μηδαμινή.

Είσοδοι – έξοδοι και τρέξιμο των προγραμμάτων:

Το σύνολο δεδομένων εκπαίδευσης και δοκιμής παρήχθησαν μέσω του προγράμματος MakeData1.java το οποίο παράγει ένα αρχείο s1data.csv το οποίο έχει 8000 (4000 train set, 4000 test set) σημεία x1,x2 και την αντίστοιχη κατηγορία που ανήκουν.

Τα προγράμματα είναι τα Ex1P1.java, Ex1P2.java.

Το compile γίνεται με την εντολή javac ExP1.java / javac ExP2.java και η εκτέλεση των προγραμμάτων με java ExP1 και java ExP2 αντίστοιχα.

Στις main τους δημιουργούνται και εκτελούνται τα νευρωνικά δίκτυα μπαίνουν ως παράμετροι διάφορες τιμές όπως ο αριθμός των νευρώνων των κρυμμένων επιπέδων, οι συναρτήσεις ενεργοποίησης, ο ρυθμός μάθησης, το κατώφλι του λάθους και το batch.

Ενδεικτικά:

```
P1 nn1 = new P1(2, // δύο είσοδοι x1,x2
```

```
4, // τέσσερις έξοδοι για τις κατηγορίες
```

```
9, // αριθμός νευρώνων στο 1ο κρυμμένο επίπεδο
```

```
9, // αριθμός νευρώνων στο 2ο κρυμμένο επίπεδο
```

```
"tanh","tanh","logistic"); // συναρτήσεις ενεργοποίησης στα 2 κρυμμένα επίπεδα και στο  
επίπεδο εξόδου
```

Ή

```
P2 nn1 = new P2(2,4,9,9,9,"tanh","tanh","tanh","tanh"); // παρόμοια με το P1 αλλά με τρία  
κρυμμένα επίπεδα
```

```
// το ίδιο και για τους δύο τύπους νευρωνικών δικτύων παρακάτω
```

```
nn1.loadData("s1data.csv"); // φορτώνονται τα δεδομένα
```

```
nn1.learningRate = 0.00005; // ρυθμός μάθησης
```

```
nn1.threshold = 0.001; // difference of error between two consecutive epochs
```

```
nn1.train(10); // εκπαίδευση με παράμετρο το B
```

```
nn1.test(); // εξέταση της γενίκευσης με το test set
```

Μετά το πέρας της εκπαίδευσης εμφανίζονται κάποια στατιστικά όπως ο χρόνος εκπαίδευσης, οι τιμές των παραμέτρων, η λίστα με τα σφάλματα (για πιθανό plotting). Μετά την εξέταση της γενίκευσης στο test set τυπώνεται η γενικευτική ικανότητα μέσω του ποσοστού σωστών προβλέψεων.

Στο τέλος υπάρχει η επιλογή να σωθούν τα στατιστικά σε ένα αρχείο P1_statistics για το 1^ο νευρωνικό δίκτυο και P2_statistics αντίστοιχα για το 2^ο για περαιτέρω σύγκριση. Επίσης, αποθηκεύεται η λίστα με τις προβλέψεις του testing για πιθανό plotting.

Πίνακας πειραμάτων των νευρωνικών δικτύων:

Στήλη	H1 neuron	H1 act. fun.	H2 neuron	H2 act. fun.	H3 neuron	H3 act. fun.	Output act. fun.	B	Learning Rate	Error Threshold	Training Time	Test correct
P1	3	tanh	3	tanh	-	-	tanh	100	0.00001	0.001	175,926 s	0.679
P1	3	relu	3	tanh	-	-	logistic	100	0.00001	0.001	28,263 s	0.53125
P1	3	tanh	3	relu	-	-	logistic	100	0.00001	0.001	12,995 s	0.44325
P1	3	relu	3	relu	-	-	logistic	100	0.00001	0.0001	19,631 s	0.34725
P1	7	tanh	4	tanh	-	-	tanh	100	0.000001	0.001	317,574 s	0.76775
P1	7	relu	4	tanh	-	-	logistic	100	0.00001	0.001	22,875 s	0.65675
P1	7	tanh	4	relu	-	-	logistic	100	0.00001	0.001	29,666 s	0.81675
P1	7	relu	4	relu	-	-	logistic	100	0.00001	0.001	18,085 s	0.27575
P1	4	tanh	9	tanh	-	-	tanh	100	0.00001	0.05	26,367 s	0.68
P1	4	relu	9	tanh	-	-	logistic	100	0.00001	0.005	50,479 s	0.55075
P1	4	tanh	9	relu	-	-	logistic	100	0.00001	0.0001	48,572 s	0.6345
P1	4	relu	9	relu	-	-	logistic	100	0.00001	0.0001	26,896 s	0.48125
P1	9	tanh	9	tanh	-	-	tanh	100	0.00001	0.001	78,490 s	0.89275
P1	9	relu	9	tanh	-	-	logistic	100	0.00001	0.001	35,889 s	0.802
P1	9	tanh	9	relu	-	-	logistic	100	0.00001	0.001	42,003 s	0.74475
P1	9	relu	9	relu	-	-	logistic	100	0.00001	0.001	32,723 s	0.60325
P1	3	tanh	3	tanh	-	-	tanh	10	0.00001	0.001	17,120 s	0.5445
P1	3	relu	3	tanh	-	-	logistic	10	0.00001	0.001	273,194 s	0.511
P1	3	tanh	3	relu	-	-	logistic	10	0.00001	0.001	70,194 s	0.50875
P1	3	relu	3	relu	-	-	logistic	10	0.00001	0.00001	37,661 s	0.28425
P1	7	tanh	4	tanh	-	-	tanh	10	0.00001	0.0001	297,926 s	0.80775
P1	7	relu	4	tanh	-	-	logistic	10	0.00001	0.0001	89,154 s	0.47025
P1	7	tanh	4	relu	-	-	logistic	10	0.00001	0.0001	113,716 s	0.7145
P1	7	relu	4	relu	-	-	logistic	10	0.00001	0.0001	117,5 s	0.58025
P1	4	tanh	9	tanh	-	-	tanh	10	0.00001	0.001	206,664 s	0.538
P1	4	relu	9	tanh	-	-	logistic	10	0.00001	0.0001	328,060 s	0.54225
P1	4	tanh	9	relu	-	-	logistic	10	0.00001	0.001	226,247 s	0.67025
P1	4	relu	9	relu	-	-	logistic	10	0.00001	0.001	58,593 s	0.49675
P1	9	tanh	9	tanh	-	-	tanh	10	0.00005	0.001	61,902 s	0.86175
P1	9	tanh	9	tanh	-	-	logistic	10	0.00005	0.001	230,646 s	0.97375
P1	9	relu	9	tanh	-	-	logistic	10	0.00001	0.001	138,578 s	0.644
P1	9	tanh	9	relu	-	-	logistic	10	0.00001	0.001	438,919 s	0.82875
P1	9	relu	9	relu	-	-	logistic	10	0.00001	0.001	77,464 s	0.51275
	H1 neurons	H1 act. fun.	H2 neurons	H2 act. fun.	H3 neurons	H3 act. fun.	Output act. fun.	B	Learning Rate	Error Threshold	Training Time	Test correct
P2	3	tanh	7	tanh	7	tanh	logistic	10	0.00001	0.001	310,549 s	0.75375
P2	7	tanh	7	tanh	7	tanh	logistic	10	0.00001	0.001	412,225 s	0.931
P2	4	tanh	4	tanh	4	tanh	logistic	10	0.00001	0.001	64,866 s	0.51875
P2	4	tanh	7	tanh	4	tanh	logistic	10	0.00001	0.001	250,504 s	0.68525
P2	4	tanh	4	tanh	7	tanh	logistic	10	0.00001	0.005	380,489 s	0.72775
P2	4	tanh	4	tanh	4	tanh	logistic	10	0.00001	0.005	389,520 s	0.92125
P2	7	tanh	7	tanh	3	tanh	logistic	10	0.00001	0.005	357,957 s	0.95275
P2	9	tanh	9	tanh	9	tanh	logistic	10	0.00001	0.005	888,249 s	0.95075
P2	11	tanh	11	tanh	11	tanh	logistic	10	0.00001	0.005	145,559 s	0.95625
P2	9	tanh	6	tanh	6	tanh	logistic	10	0.00001	0.01	39,017 s	0.883
P2	6	tanh	6	tanh	6	tanh	logistic	10	0.00001	0.01	47,205 s	0.92975
P2	11	tanh	11	tanh	11	tanh	logistic	100	0.00001	0.005	145,559 s	0.95625
P2	6	relu	6	tanh	6	tanh	logistic	100	0.00001	0.01	34,546 s	0.6565
P2	6	relu	6	tanh	6	relu	logistic	100	0.00001	0.01	38,689 s	0.438
P2	6	relu	6	relu	6	relu	logistic	100	0.00001	0.01	53,126 s	0.39075
P2	6	tanh	6	relu	6	relu	logistic	100	0.00001	0.01	27,534 s	0.57325
P2	6	tanh	6	tanh	6	relu	logistic	100	0.00001	0.01	32,989 s	0.8135
P2	6	tanh	6	relu	6	tanh	logistic	100	0.00001	0.01	35,485 s	0.708
P2	6	tanh	6	tanh	6	tanh	logistic	100	0.00001	0.01	64,930 s	0.918
P2	9	tanh	9	tanh	5	relu	logistic	100	0.00001	0.01	46,532 s	0.89725
P2	9	relu	9	tanh	5	tanh	logistic	100	0.00001	0.01	44,059 s	0.8295
P2	9	tanh	9	tanh	5	tanh	logistic	100	0.00001	0.01	67,636 s	0.8885
P2	9	tanh	9	tanh	5	tanh	logistic	100	0.000005	0.001	152,158 s	0.9645
P2	9	tanh	9	tanh	5	tanh	tanh	100	0.000005	0.001	121,568 s	0.91325

Αποτύπωση του συνόλου εκπαίδευσης του καλύτερου νευρωνικού δικτύου (97.375% σωστά):

