

Υπολογιστική Νοημοσύνη
-Τέταρτη Εργασία-

Αλέξανδρος Πετρίδης

Τελευταία ενημέρωση: 7 Μαρτίου 2022

Περιεχόμενα

| | | |
|----------|---------------------------------------|----------|
| 1 | Στόχος | 3 |
| 2 | Απλή εφαρμογή RBF δικτύου | 3 |
| 2.1 | Αποτελέσματα και σχολιασμός | 3 |
| 3 | Fine tuning δικτύου | 3 |

Κατάλογος Σχημάτων

| | | |
|---|---|---|
| 1 | Καμπύλες εκμάθησης με αριθμό νευρώνων του ενδιάμεσου (RBF) κρυφού στρώματος (10%, 50%, 90%) του πλήθους των δεδομένων εκπαίδευσης | 3 |
| 2 | Καμπύλη εκμάθησης | 4 |

Κατάλογος Πινάκων

| | | |
|---|--------------------------------|---|
| 1 | Μετρικές αξιολόγησης | 3 |
| 2 | Βέλτιστες Παράμετροι | 4 |

1 Στόχος

Στόχος είναι η επίλυση ενός προβλήματος παλινδρόμησης με χρήση RBF δικτύου. Το dataset που επιλέχθηκε ήταν το Boston housing dataset το οποίο περιλαμβάνει 506 δείγματα με 14 χαρακτηριστικά το καθένα, για την πρόβλεψη της τιμής ακινήτων στην περιοχή της Βοστώνης.

2 Απλή εφαρμογή RBF δικτύου

Αρχικά, διαχωρίστηκαν τα δεδομένα σε υποσύνολα εκπαίδευσης και ελέγχου, με αναλογία 75% - 25%. Έπειτα σχεδιάστηκαν τρία απλά RBF δίκτυα, με αριθμό νευρώνων του ενδιάμεσου (RBF) κρυφού στρώματος (10%, 50%, 90%) του πλήθους των δεδομένων εκπαίδευσης. Η εκπαίδευση έγινε για 100 εποχές, με ρυθμό εκπαίδευσης $l_r = 0.001$, ενώ το 20% των δεδομένων εκπαίδευσης παρακρατήθηκε για επικύρωση. Το πλήθος των νευρώνων για το στρώμα εξόδου θεωρήθηκε 128. Όμως στο τέλος χρησιμοποιήθηκε άλλο ένα στρώμα με μια έξοδο για να μπορέσουν να υπολογιστούν οι παρακάτω μετρητές που ζητούνται.

2.1 Αποτελέσματα και σχολιασμός

Ακολουθούν οι καμπύλες εκμάθησης για training και validation δεδομένα.



Σχήμα 1: Καμπύλες εκμάθησης με αριθμό νευρώνων του ενδιάμεσου (RBF) κρυφού στρώματος (10%, 50%, 90%) του πλήθους των δεδομένων εκπαίδευσης

Ακόμα οι μετρίκες αξιολόγησης που υπολογίστηκαν είναι οι παρακάτω.

| Μετρίκες | R^2 | $RMSE$ |
|----------------|-------|--------|
| Μοντέλο με 10% | 0.376 | 6.881 |
| Μοντέλο με 50% | 0.418 | 6.646 |
| Μοντέλο με 90% | 0.464 | 6.375 |

Πίνακας 1: Μετρίκες αξιολόγησης

Από την θεωρία είναι γνωστό πως το $RMSE$ είναι το ποσοστό σφάλματος κατά την τετραγωνική ρίζα του MSE , όπου το MSE αντιπροσωπεύει τη διαφορά μεταξύ των αρχικών και των προβλεπόμενων τιμών που εξάγονται με το τετράγωνο της μέσης διαφοράς στο σύνολο δεδομένων. Όσο μικρότερη η τιμή, τόσο καλύτερο είναι το μοντέλο. Συγκεκριμένα $RMSE =$

$$\sqrt{MSE} = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^N (y_i - \hat{y})^2}.$$

Ακόμα το $R^2 = 1 - \frac{\sum (y_i - \hat{y})^2}{\sum (y_i - \bar{y})^2}$ αντιπροσωπεύει το συντελεστή του πόσο καλά ταιριάζουν οι προβλεπόμενες τιμές σε σύγκριση με τις αρχικές. Σε αυτήν την περίπτωση όσο μεγαλύτερη είναι η τιμή, τόσο καλύτερο είναι το μοντέλο.

Από τις παραπάνω τιμές των μετρίκων παρατηρείται πως δεν υπάρχουν μεγάλες αποκλίσεις στα μοντέλα που δημιουργήθηκαν και εκπαιδεύτηκαν. Ένα θετικό σημείο είναι πως το R^2 δεν είναι αρνητικός αριθμός, ενώ η μετρική $RMSE$ δεν έχει πολύ μεγάλες τιμές. Μπορούμε ακόμα να ισχυριστούμε πως το καλύτερο μοντέλο είναι αυτό με τους περισσότερους νευρώνες στο επίπεδο RBF, όμως ακόμα παρατηρούμε ότι οι διαφορές είναι πάρα πολύ μικρές.

3 Fine tuning δικτύου

Στο συγκεκριμένο κομμάτι της εργασίας, σκοπός είναι η εύρεση των βέλτιστων τιμών για μερικές υπερπαραμέτρους του δικτύου, και η τελική εκπαίδευση και αξιολόγηση ενός μοντέλου με βάση της επιλεγμένες τιμές.

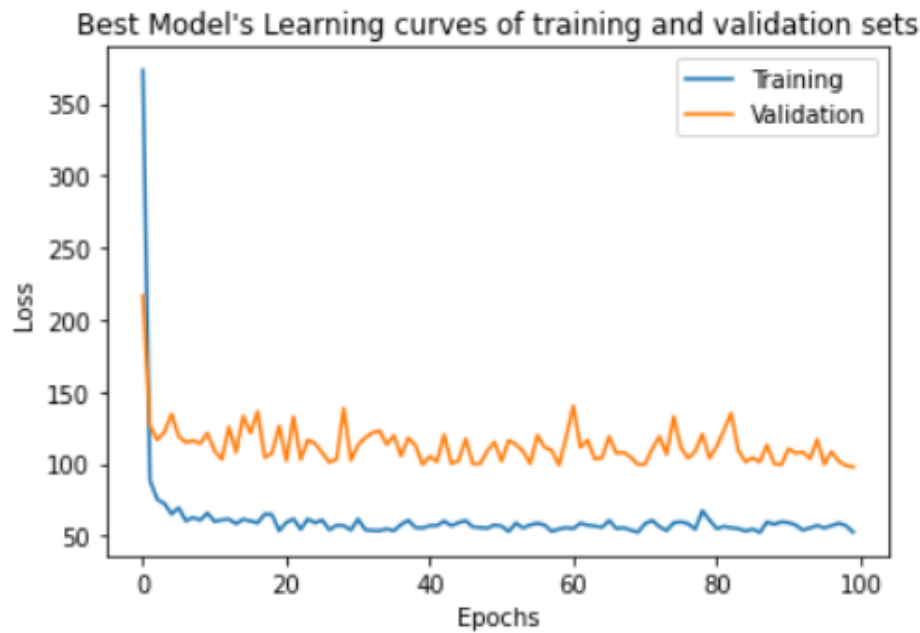
Το RBF δίκτυο που εκπαιδεύτηκε στο παρόν κομμάτι είχε την ίδια αρχιτεκτονική με τα παραπάνω μοντέλα (ένα RBF κρυφό στρώμα και ένα γραμμικό στρώμα εξόδου), με την επιπλέον προσθήκη dropout κανονικοποίησης για τους νευρώνες του στρώματος εξόδου. Οι υπερπαράμετροι του δικτύου προς εξέταση, καθώς και το εύρος αναζήτησης για κάθε παράμετρο, ήταν:

- αριθμός νευρώνων RBF στρώματος: $n_{h1} \in 5\%, 15\%, 30\%, 50\%$ του πλήθους δεδομένων εκπαίδευσης
- αριθμός νευρώνων δεύτερου κρυφού στρώματος: $n_{h2} \in 32, 64, 128, 256$
- dropout πιθανότητα $p \in 0.2, 0.35, 0.5$

Μετά την εκπαίδευση του βέλτιστου μοντέλου βρέθηκε πως είχε $R^2 = 0.303$ και $RMSE = 7.272$. Τιμές όχι πολύ διαφορετικές από τα μοντέλα της προηγούμενης ενότητας. Ακολουθεί η καμπύλη εκμάθησης του Fine Tuning δικτύου. Το οποίο είχε σαν βέλτιστες παραμέτρους τις παρακάτω:

| | |
|---|-----|
| Βέλτιστος αριθμός νευρώνων πρώτου κρυφού RBF επιπέδου | 19 |
| Βέλτιστος αριθμός νευρώνων δεύτερου κρυφού επιπέδου | 128 |
| Βέλτιστη dropout πιθανότητα | 0.5 |

Πίνακας 2: Βέλτιστες Παράμετροι



Σχήμα 2: Καμπύλη εκμάθησης