

Εργασία 4: Υπολογιστική Νοημοσύνη

Επίλυση προβλήματος παλινδρόμησης με χρήση RBF δικτύου

Στόχος της εργασίας είναι ο πειραματισμός πάνω σε μια απλή αρχιτεκτονική RBF για την επίλυση ενός απλού προβλήματος παλινδρόμησης. Επιλέγεται το Boston housing dataset το οποίο περιλαμβάνει 506 δείγματα με 14 χαρακτηριστικά το καθένα, για την πρόβλεψη της τιμής ακινήτων στην περιοχή της Βοστώνης.

1 Επισκόπηση RBF δικτύων

Τα RBF δίκτυα αποτελούν ένα συγκεκριμένο τύπο νευρωνικού δικτύου με 3 συνολικά στρώματα:

- Το πρώτο στρώμα το οποίο δέχεται την είσοδο του μοντέλου
- Το κρυφό στρώμα το οποίο μετασχηματίζει την εκάστοτε είσοδο με μια Radial Basis Function

$$f(\mathbf{x}) = \frac{e^{-\|\mathbf{x}-\mathbf{c}_i\|^2}}{2\sigma_i^2}$$

, όπου c_i και σ_i το κέντρο και η διασπορά της συνάρτησης του i -στού νευρώνα

- Το στρώμα εξόδου το οποίο εκτελεί ένα γραμμικό μετασχηματισμό πάνω στην έξοδο του κρυφού στρώματος, για να παράξει την τελική έξοδο του μοντέλου

$$\hat{y} = \mathbf{W}^T f(\mathbf{x}) + \mathbf{b}$$

, όπου \mathbf{W} και \mathbf{b} τα συναπτικά βάρη και τα κατώφλια του στρώματος εξόδου.

Ένα RBF δίκτυο εκπαιδεύεται σε δύο φάσεις:

1. Αρχικά, επιλέγονται τιμές για τα \mathbf{c} και σ , συνήθως με κάποιον αλγόριθμο ομαδοποίησης δεδομένων

2. Με τις παραπάνω επιλεγμένες τιμές, εκπαιδεύεται το στρώμα εξόδου του δικτύου με τις γνωστές μεθόδους εκπαίδευσης νευρωνικών δικτύων, για να υπολογιστούν τα \mathbf{W} και \mathbf{b} .

2 Απλή εφαρμογή σε RBF δίκτυο

Αρχικά, να διαχωριστούν τα δεδομένα σε υποσύνολα εκπαίδευσης και ελέγχου, με αναλογία 75% – 25%. Να σχεδιαστούν και να εκπαιδευτούν τρία απλά RBF δίκτυα, με αριθμό νευρώνων του ενδιαμέσου (RBF) κρυφού στρώματος (10%, 50%, 90%) του πλήθους των δεδομένων εκπαίδευσης. Ο υπολογισμός των κεντρών \mathbf{c} να γίνει με χρήση του KMeans αλγορίθμου ομαδοποίησης, ενώ η διασπορά κάθε νευρώνα (για κάθε μια από τις τρεις περιπτώσεις) να τεθεί στην τιμή

$$\sigma = \frac{d_{max}}{\sqrt{2P}}$$

, όπου d_{max} η μέγιστη Για κάθε ένα μοντέλο, μετά τον υπολογισμό των (\mathbf{c}, σ) να χρησιμοποιηθεί ο SGD βελτιστοποιητής για την εκπαίδευση του στρώματος εξόδου \mathbf{W}, \mathbf{b} . Η εκπαίδευση να γίνει για 100 εποχές, με ρυθμό εκμάθησης $lr = 0.001$. Το πλήθος των νευρώνων για το στρώμα εξόδου να θεωρηθεί 128.

Για τις δύο παραπάνω φάσεις εκπαίδευσης, να παρακρατηθεί το 20% των δεδομένων εκπαίδευσης για επικύρωση. Μετά τη λήξη της διαδικασίας, ζητούνται για κάθε μοντέλο τα εξής:

1. Καμπύλες εκμάθησης για training και validation δεδομένα
2. Οι μετρικές αξιολόγησης R^2 και $RMSE$
3. Σχολιασμός των αποτελεσμάτων όσον αφορά κυρίως τον αριθμό νευρώνων στο RBF στρώμα.

3 Fine tuning RBF δικτύου

Στο συγκεκριμένο κομμάτι της εργασίας, σκοπός είναι η εκτέλεση εύρεση των βέλτιστων τιμών για μερικές υπερπαραμέτρους του δικτύου, και η τελική εκπαίδευση και αξιολόγηση ενός μοντέλου με βάση τις επιλεγμένες τιμές. Το RBF δίκτυο που θα εξεταστεί στο παρόν κομμάτι θα έχει 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, θα είναι:

1. αριθμός νευρώνων RBF στρώματος: $n_{h_1} \in \{5\%, 15\%, 30\%, 50\%\}$ του πλήθους δεδομένων εκπαίδευσης
2. αριθμός νευρώνων δεύτερου κρυφού στρώματος: $n_{h_2} \in \{32, 64, 128, 256\}$
3. dropout πιθανότητα $p \in \{0.2, 0.35, 0.5\}$

Η αρχικοποίηση των διασπορών των νευρώνων του RBF στρώματος να γίνει όπως και στο πρώτο τμήμα. Η εκπαίδευση του συνολικού δικτύου να γίνεται σε δύο φάσεις, όπως περιγράφηκε παραπάνω, με χρήση του Kmeans για το πρώτο στρώμα και SGD optimizer ($lr = 0.001$, 100 εποχές) για το δεύτερο στρώμα. Ως συνάρτηση κόστους θεωρείται το μέσο τετραγωνικό σφάλμα και ως μετρική αξιολόγησης το *RMSE*.

Η διαδικασία εύρεσης των βέλτιστων τιμών των παραμέτρων μπορεί να γίνει είτε με χρήση grid search και 5-fold Cross Validation, είτε με τη χρήση ενός διαθέσιμου framework (όπως πχ ο keras-tuner). Σε κάθε περίπτωση, 20% των δεδομένων εκπαίδευσης παρακρατείται για επικύρωση.