

Υπολογιστική Νοημοσύνη

Εργασία 4

Επίλυση προβλήματος παλινδρόμησης με χρήση RBF δικτύου

Ονοματεπώνυμο: Σιδηρόπουλος Λεωνίδας

AEM: 9818

email: leonsidi@ece.auth.gr

Περίοδος: Φεβρουάριος 2023

Απλή εφαρμογή σε RBF δίκτυο

Για την απλή εφαρμογή σε RBF δίκτυο, χρησιμοποιήθηκαν:

- Ένα στρώμα εισόδου του μοντέλου (RBF Layer).
- Ένα κρυφό στρώμα που αποτελείται από 128 νευρώνες (DenseLayer).
- Ένα στρώμα εξόδου που αποτελείται από 1 νευρώνα (OutLayer).
- Το Boston Housing dataset της βιβλιοθήκης keras.

Για την υλοποίηση του δικτύου, χρησιμοποιήθηκαν 3 διαφορετικά μοντέλα, ως προς των αριθμό των νευρώνων που περιέχει το καθένα.

Για την εκπαίδευση, γίνεται χρήση του SGD optimizer, για 100 εποχές (epochs) και με ρυθμό εκμάθησης (learning_rate) η = 0,001. Επίσης, έχουμε ορίσει το batch size = 16.

Για τον υπολογισμό του R^2 (Coefficient of Determination) χρησιμοποιήθηκε η παρακάτω μαθηματική φόρμουλα:

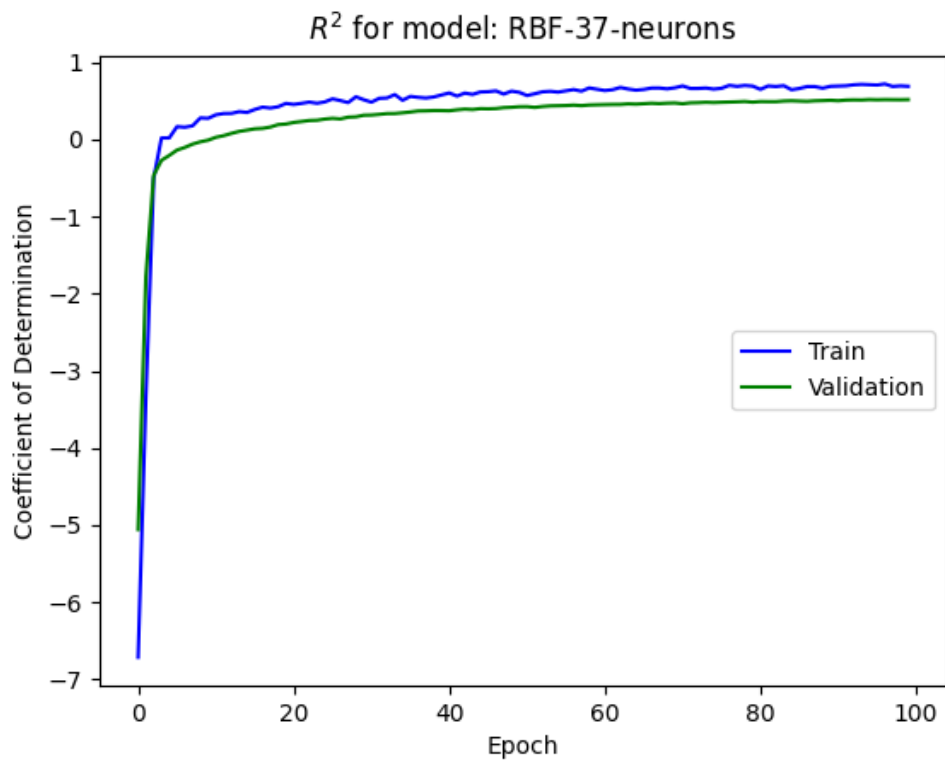
$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

όπου:

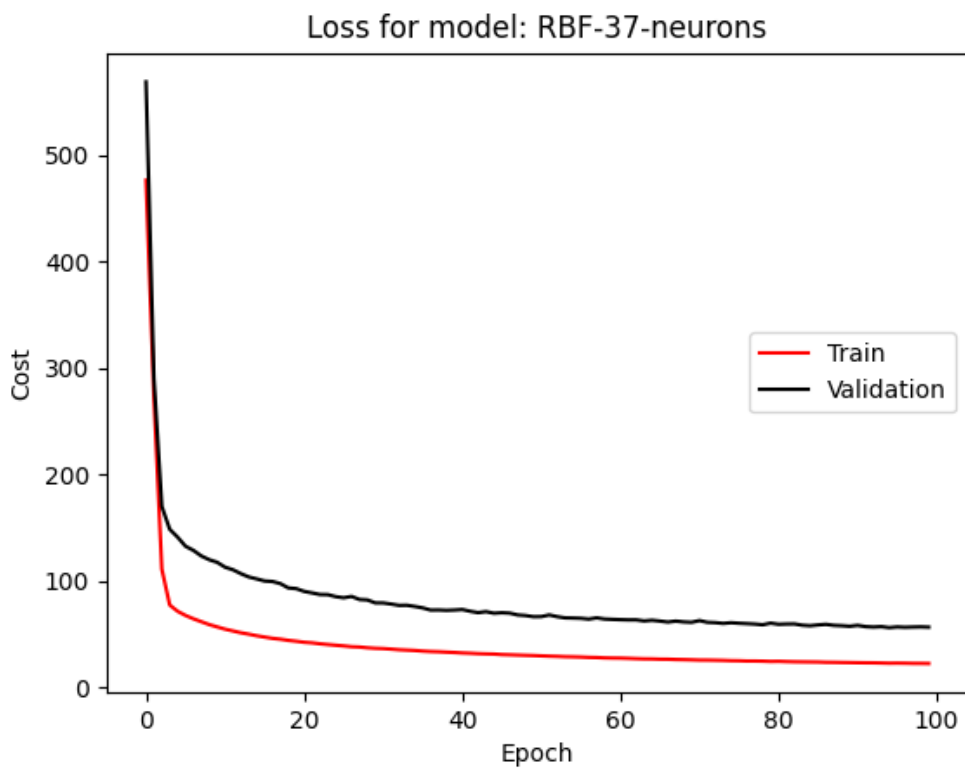
- SS_{res} είναι το άθροισμα τετραγώνων υπολοίπων
- SS_{tot} είναι το συνολικό άθροισμα τετραγώνων

Μοντέλο 1

Το μοντέλο 1 αποτελείται από 37 νευρώνες (10% του πλήθους των δεδομένων εκπαίδευσης). Λόγω του μικρού πλήθους νευρώνων του συγκεκριμένου μοντέλου, αναμένουμε να έχει ταχύτερη εκπαίδευση σε σύγκριση με τα υπόλοιπα 2 μοντέλα. Στα παρακάτω διαγράμματα, παρουσιάζονται οι καμπύλες εκπαίδευσης (Train) και επαλήθευσης (Validation) για το R^2 και τις απώλειες.



Διάγραμμα R^2 για το μοντέλο νευρωνικού δικτύου με 37 νευρώνες

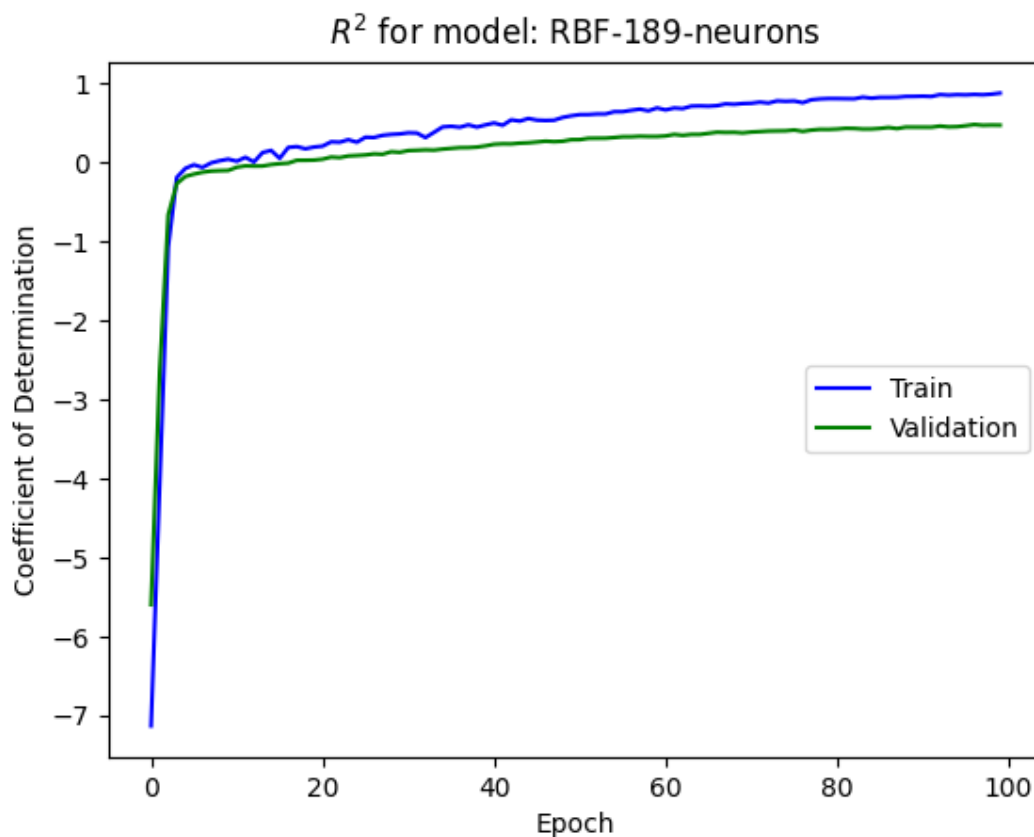


Διάγραμμα απωλειών για το μοντέλο νευρωνικού δικτύου με 37 νευρώνες

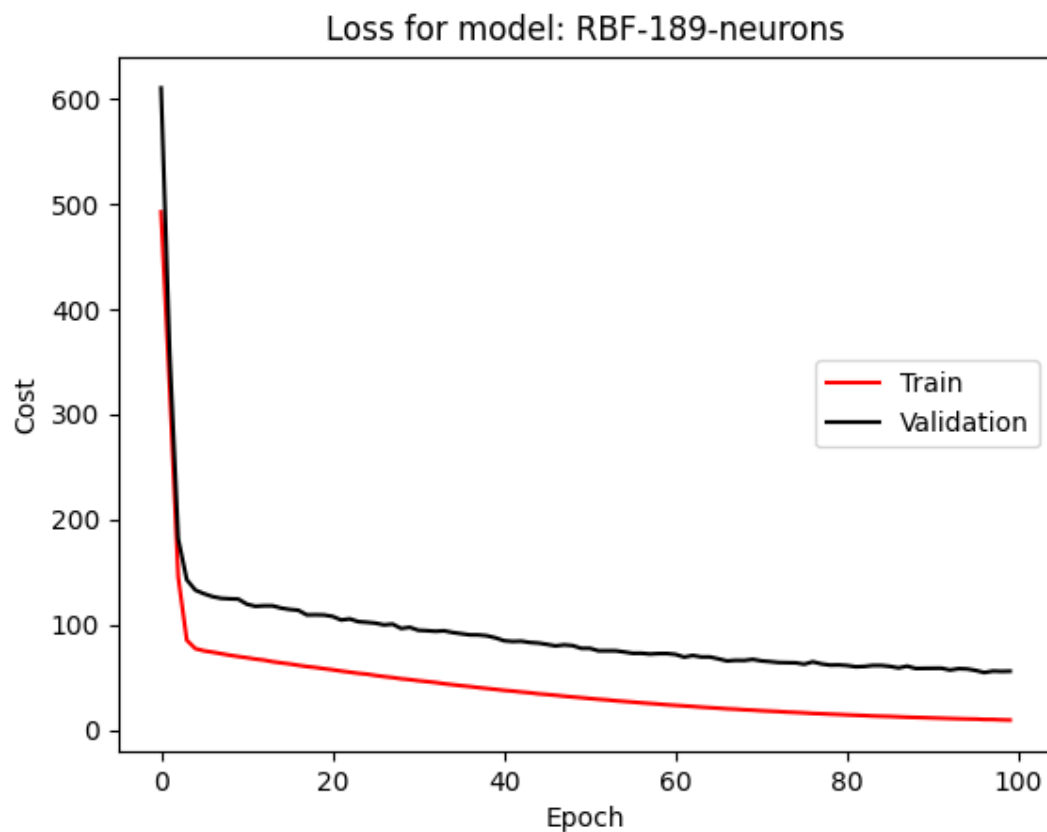
Παρατηρούμε ότι οι μεγάλες τιμές κόστους, στο διάγραμμα απωλειών, και RMSE υποδηλώνουν ότι η εκπαίδευση του μοντέλου δεν ήταν πολύ αποτελεσματική. Ωστόσο, παρατηρούμε στη συνέχεια, με την αύξηση των εποχών, ότι το κόστος μειώνεται κατα πολύ και το R^2 αυξάνεται.

Μοντέλο 2

Το μοντέλο 2 αποτελείται από 189 νευρώνες (50% του πλήθους των δεδομένων εκπαίδευσης). Αναμένουμε πιο αργή εκπαίδευση σε σύγκριση με το μοντέλο 1, λόγω του μεγάλου αριθμού νευρώνων στο μοντέλο 2. Τα διαγράμματα για το R^2 και τις απώλειες είναι τα παρακάτω:



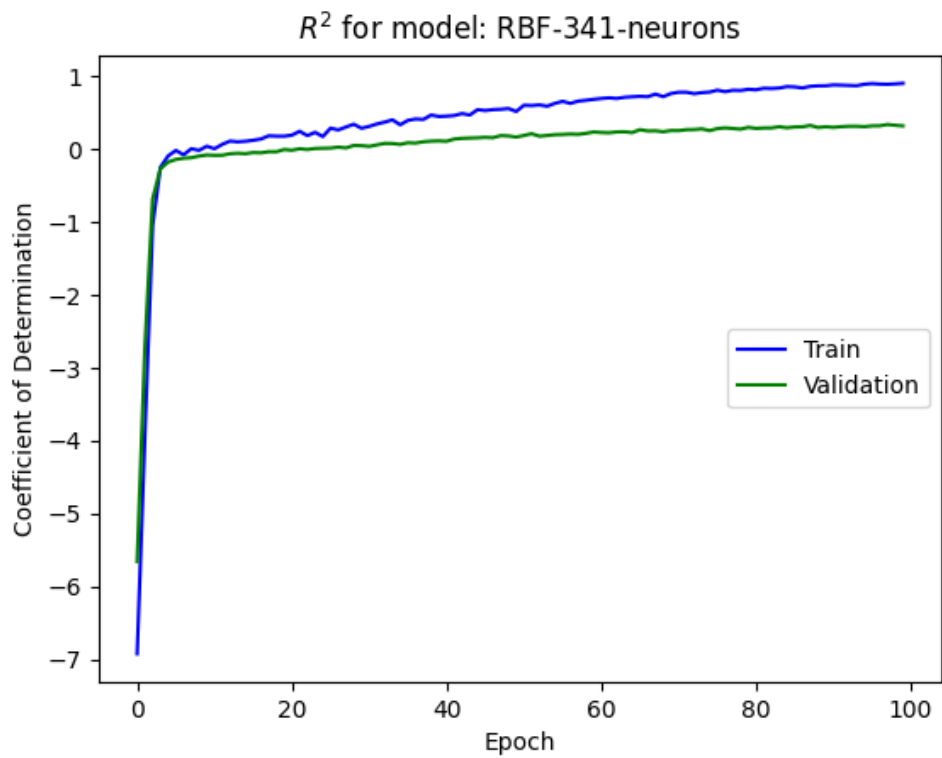
Διάγραμμα R^2 για το μοντέλο νευρωνικού δικτύου με 189 νευρώνες



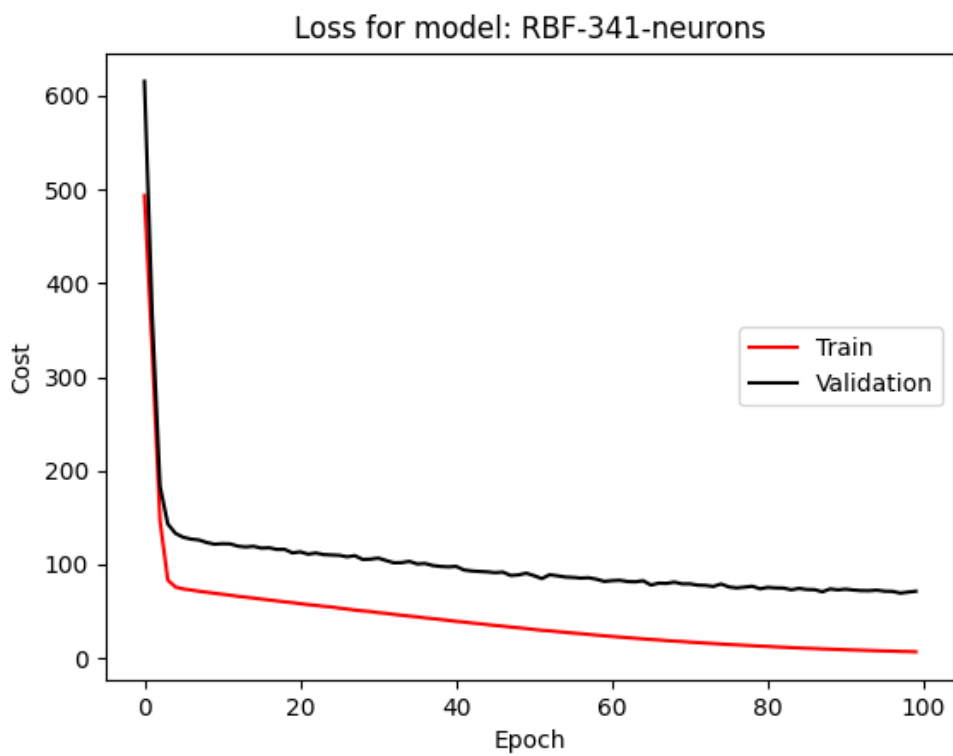
Διάγραμμα απωλειών για το μοντέλο νευρωνικού δικτύου με 189 νευρώνες

Μοντέλο 3

Το μοντέλο 3 αποτελείται από 341 νευρώνες (90% του πλήθους των δεδομένων εκπαίδευσης). Τα διαγράμματα για το R^2 και τις απώλειες είναι τα παρακάτω:



Διάγραμμα R^2 για το μοντέλο νευρωνικού δικτύου με 341 νευρώνες



Διάγραμμα απωλειών για το μοντέλο νευρωνικού δικτύου με 341 νευρώνες

Παρατηρήσεις και Αποτελέσματα

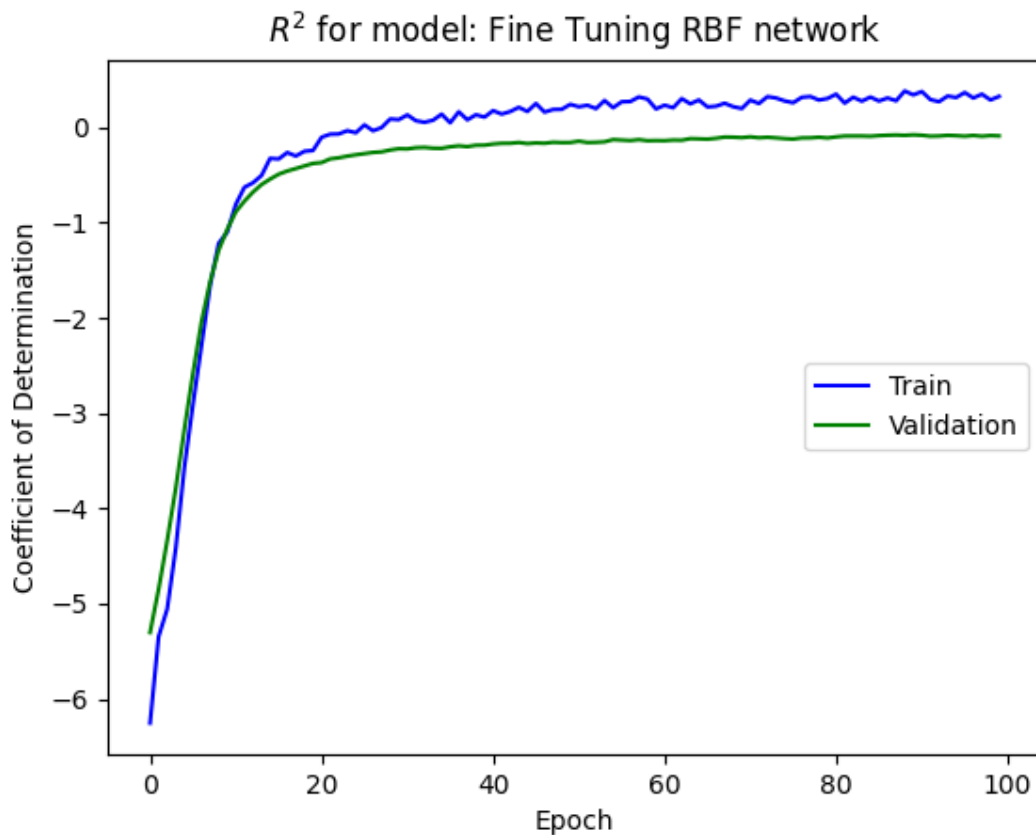
Τα αποτελέσματα της εκπαίδευσης και των τριών μοντέλων παρουσιάζονται στον παρακάτω πίνακα:

Μοντέλο	R^2	RMSE	Validation R^2	Validation RMSE
1	0.68629	4.77809	0.51408	7.54674
2	0.88127	3.04224	0.47539	7.46781
3	0.90207	2.60724	0.31807	8.45049

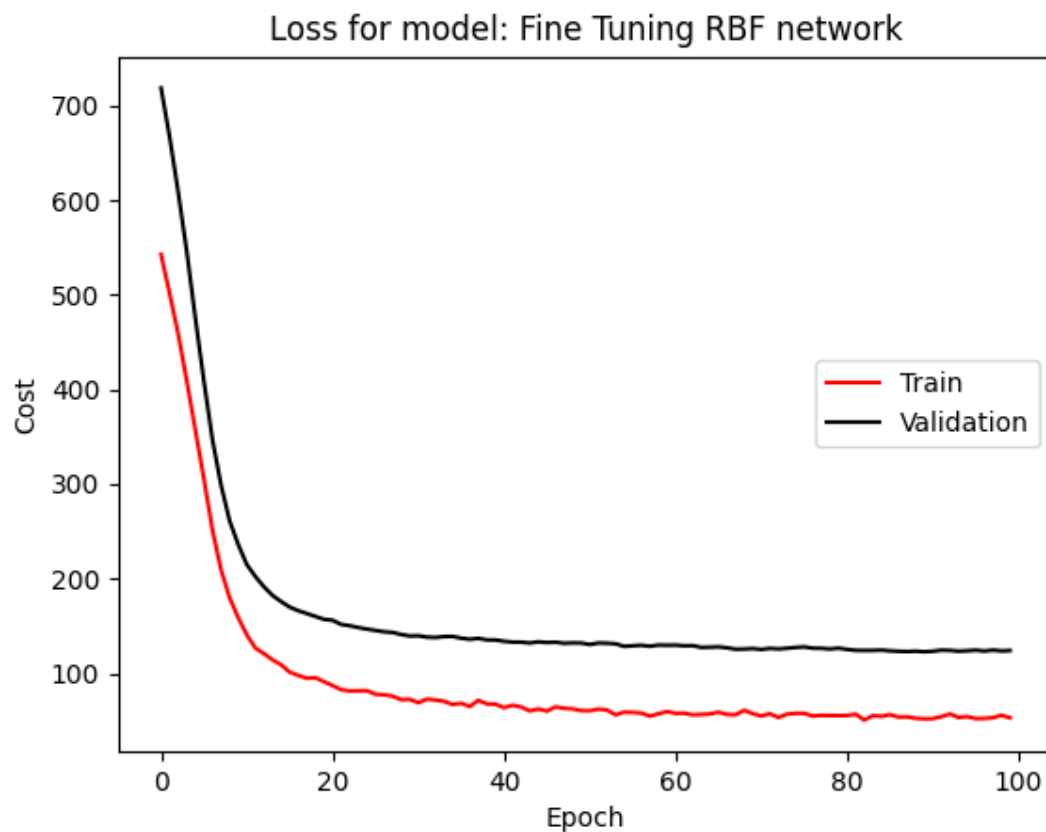
Από τα παραπάνω αποτελέσματα παρατηρούμε ότι το μοντέλο 1 έχει το μεγαλύτερο RMSE και άρα δεν είναι τόσο αποτελεσματικό όσο τα μοντέλα 2 και 3 κάτι που επιβεβαιώνεται και από τις τιμές για το R^2 . Τα μοντέλα 2 και 3 έχουν τις τιμές R^2 και RMSE αρκετά κοντά μεταξύ τους. Αυτό σημαίνει ότι, από ένα σημείο και μετά η αύξηση των νευρώνων του RBF δικτύου δεν έχει σημαντική επίδραση στη βελτίωση της εκπαίδευσης του μοντέλου.

Fine tuning RBF δικτύου

Για την συγκεκριμένη υλοποίηση, δημιουργούμε ένα πλέγμα τιμών για τις υπερπαραμέτρους του κάθε μοντέλου και πραγματοποιούμε grid search για την εύρεση του καλύτερου μοντέλου. Η επιλογή του βέλτιστου μοντέλου γίνεται με βάση το καλύτερο RMSE. Έπειτα, για την αξιολόγηση του κάθε μοντέλου εφαρμόζουμε 5-fold Cross Validation, το οποίο μεν είναι κοστοβόρο ως προς τον χρόνο εκτέλεσης, ωστόσο έχουμε πιο εύρωστα αποτελέσματα. Για την εκπαίδευση του μοντέλου, γίνεται χρήση του SGD optimizer, για 100 εποχές (epochs), με ρυθμό εκμάθησης (learning_rate) $\eta = 0,001$ και με batch size = 128. Τα διαγράμματα για το R^2 και τις απώλειες είναι τα παρακάτω:



Διάγραμμα R^2 για το μοντέλο νευρωνικού δικτύου με tuning



Διάγραμμα απωλειών για το μοντέλο νευρωνικού δικτύου με tuning

Τα αποτελέσματα για το βέλτιστο μοντέλο συνοψίζονται στον παρακάτω πίνακα:

R^2	RMSE	Loss	RBF Layer size	Hidden Layer size	Dropout Rate
0.28539	7.34060	53.88448	0.5	64	0.2

όπου:

- RBF Layer size είναι το ποσοστό των νευρώνων στο RBF στρώμα (0.5 ή 50%).
- Hidden Layer size είναι ο αριθμός των νευρώνων στο κρυφό στρώμα
- Dropout Rate είναι η dropout πιθανότητα

Παρατηρούμε ότι, ο σημαντικότερος παράγοντας στην αξιολόγηση του μοντέλου είναι ο αριθμός των νευρώνων στον RBF layer, καθώς επιλέχτηκε η μέγιστη τιμή (0.5) από το εύρος τιμών που μας δόθηκε. Στη συνέχεια, ακολουθεί ο αριθμός των νευρώνων στον hidden layer (64) και η dropout πιθανότητα που επιλέχθηκε (0.2) είναι η μικρότερη από το εύρος τιμών που μας δόθηκε. Αυτό σημαίνει πως, πιθανότατα, δεν χρειάζεται να γίνει κάποια κανονικοποίηση του μοντέλου, διότι δεν έχουμε υπερπροσαρμογή του μοντέλου. Γενικά, παρατηρούμε ότι η απόδοση του μοντέλου δεν είναι πολύ καλή, καθώς έχουμε και πάλι υψηλές τιμές απωλειών (Loss) και RMSE.