



Προγραμματιστική άσκηση 4 - Radial Basis Function Network

Όνομα: Μιχαήλ-Παναγιώτης Μπόφος

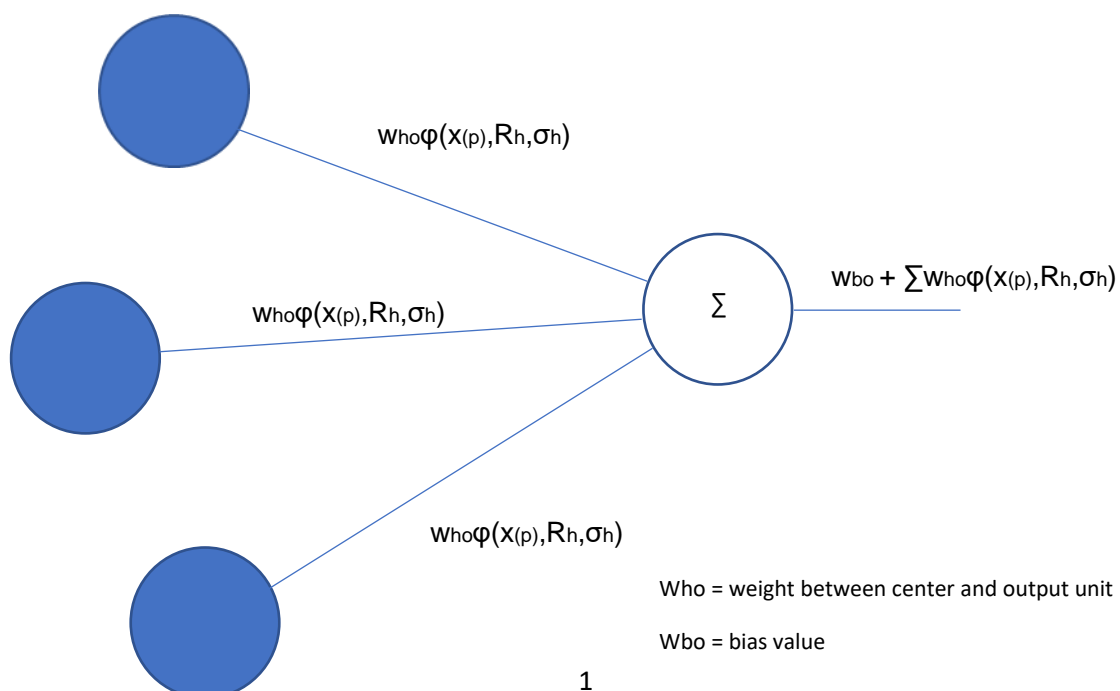
Ταυτότητα: ΑΟ082689

Δομή Δικτύου

Η δομή του δικτύου μας καθορίζεται σε μεγάλο βαθμό από τη δομή των δεδομένων μας. Ο στόχος μας είναι με βάση τα πενήντα - τρία πειραματικά δεδομένα κάθε μορίου να μπορέσουμε να υπολογίσουμε τη βιολογική του δράση. Επομένως, το αποτέλεσμα του δικτύου μας είναι μια τιμή άρα θα χρειαστούμε μόνο ένα output unit. Όσο αφορά τον αριθμό των κέντρων θα το διερευνήσουμε παρακάτω. Ωστόσο, αυτό που ξέρουμε εξ αρχής είναι ότι κάθε διάνυσμα εισόδου έχει μέγεθος πενήντα - τρία, άρα η θέση κάθε κέντρου αποτελείται και αυτή από πενήντα τρεις τιμές. Ταυτόχρονα κάθε κέντρο έχει τη δική του τιμή σ (πλάτος γκαουσιανής συνάρτησης). Οι αρχικές θέσεις, το αρχικό σ και ο αριθμός των κέντρων καθορίζονται από τον χρήστη. Σημαντική σημείωση, τα δεδομένα μας κανονικοποιήθηκαν πριν την εκπαίδευση του δικτύου μας.

Συνδεσμολογία & Λειτουργία

Το δίκτυό μας είναι πλήρως συνδεδεμένο από το επίπεδο των κέντρων (hidden layer) στο επίπεδο εξόδου (output layer). Την ένωση κάθε κέντρου με το output unit την ορίζει μία τιμή βάρους (η αρχικοποίηση της γίνεται με την `Math.random()` και η αρχική τιμή κυμαίνεται μεταξύ του -1 με το +1). Για τη μάθηση του δικτύου μας θα χρησιμοποιήσουμε τον αλγόριθμο των μεταβαλλόμενων κέντρων, άρα θα έχουμε έξι μεταβλητές (θέσεις κέντρων, πλάτος γκαουσιανής, ρυθμοί μάθησης και αριθμός επαναλήψεων). Βάσει της θεωρίας κάθε κέντρο αντικατοπτρίζει την κατανομή των δεδομένων εισόδου στο χώρο. Με τη χρήση της ευκλείδειας απόστασης μεταξύ της θέσης του και των δεδομένων εισόδου το κάθε κέντρο υπολογίζει την έξοδό του (γκαουσιανή ευκλείδεια επί βάρους κέντρου). Έπειτα, το output unit έχει ως τιμή εξόδου το άθροισμα των εξόδων των κέντρων συν μια τιμή bias.

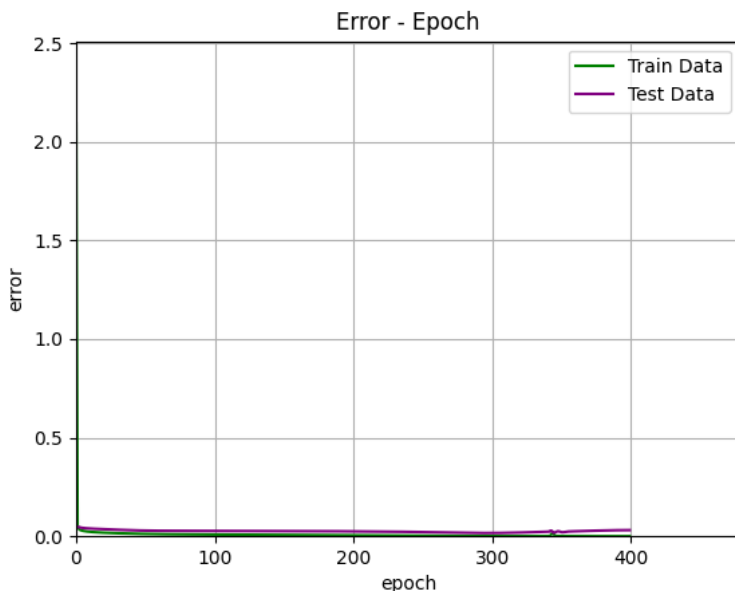


Επιλογή Μεταβλητών

Στην προκειμένη περίπτωση θέλουμε παλινδρόμηση κατά προσέγγιση και όχι παρεμβολή καθώς θέλουμε να γενικεύσουμε τη γνώση μας (πχ να υπολογίσουμε τη βιολογική δράση ενός καινούργιου μορίου). Ως εκ τούτου, σύμφωνα με τη θεωρία θέλουμε λίγα κέντρα (λιγότερα από τα δεδομένα εισόδου) και μεγάλο πλάτος γκαουσιανής (τι ορίζουμε ως μεγάλο και μικρό θα το εξερευνήσουμε αργότερα - πειραματικά). Η πειραματική μέθοδος που ακολουθήσαμε είναι απλή, δοκιμάσαμε πολλαπλούς συνδυασμούς από ρυθμούς μάθησης, πλήθος κέντρων, εποχών και πλάτους γκαουσιανών. Όσο αφορά τον αριθμό των εποχών βάσει δοκιμών καταλήξαμε ότι οι τετρακόσιες ήταν αρκετές (σε περιπτώσεις με παραπάνω εποχές κάποιες φορές παρατηρήσαμε απρόβλεπτα αποτελέσματα).

Πειραματικά Αποτελέσματα

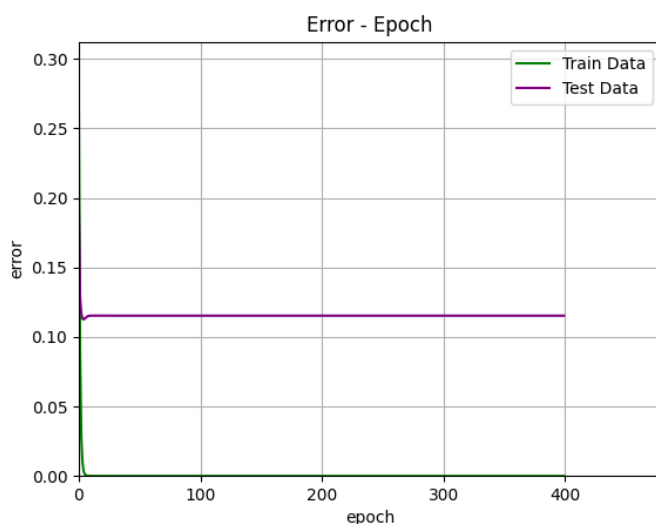
1) Είκοσι κέντρα, ένα κέντρο για κάθε γραμμή εισόδου, μεγάλο σ .



```
numHiddenLayerNeurons 20
numInputNeurons 53
numOutputNeurons 1
learningRates 0.5 0.9 0.4
sigmas 2
maxIterations 400
centresFile bigcenters.txt
trainFile train.txt
testFile test.txt
```

399 2.542505178374277E-4 0.03119752443200186

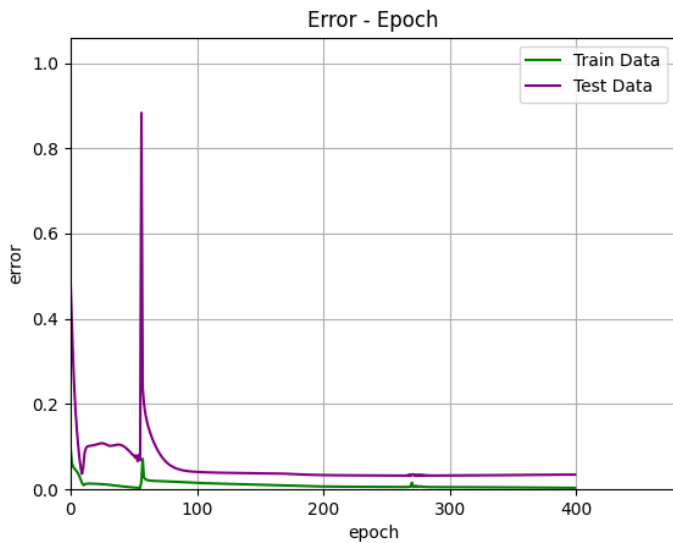
2) Είκοσι κέντρα, ένα κέντρο για κάθε γραμμή εισόδου, μικρό σ .



```
numHiddenLayerNeurons 20
numInputNeurons 53
numOutputNeurons 1
learningRates 0.5 0.9 0.4
sigmas 0.2
maxIterations 400
centresFile bigcenters.txt
trainFile train.txt
testFile test.txt
```

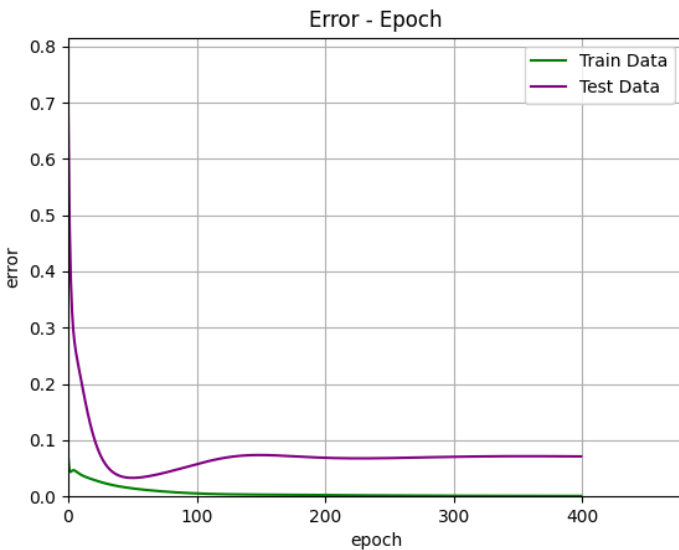
399 2.0445267214765693E-33 0.11530841494048842

3) Οκτώ κέντρα, τυχαία επιλογή κέντρων, μεγάλο σ .



```
numHiddenLayerNeurons 8
numInputNeurons 53
numOutputNeurons 1
learningRates 0.5 0.9 .4
sigmas 2
maxIterations 400
centresFile centers.txt
trainFile train.txt
testFile test.txt
```

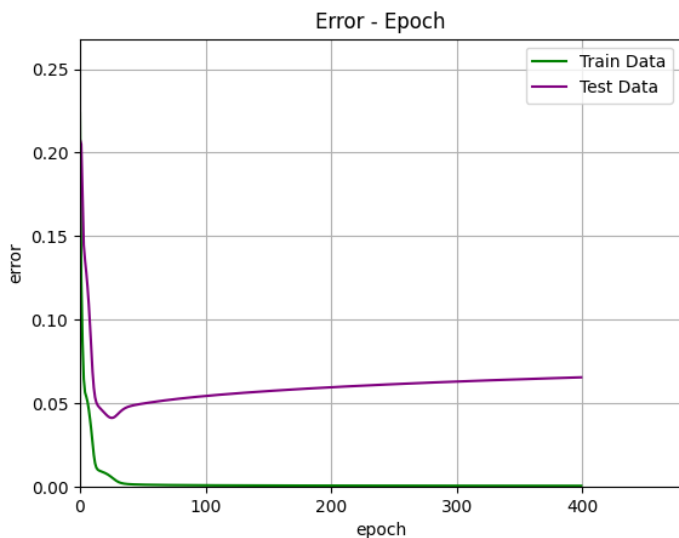
399 0.0036382772677780945 0.03474675007181288



```
numHiddenLayerNeurons 8
numInputNeurons 53
numOutputNeurons 1
learningRates 0.5 0.01 0.4
sigmas 2
maxIterations 400
centresFile centers.txt
trainFile train.txt
testFile test.txt
```

399 8.149836507851612E-4 0.07116186507984358

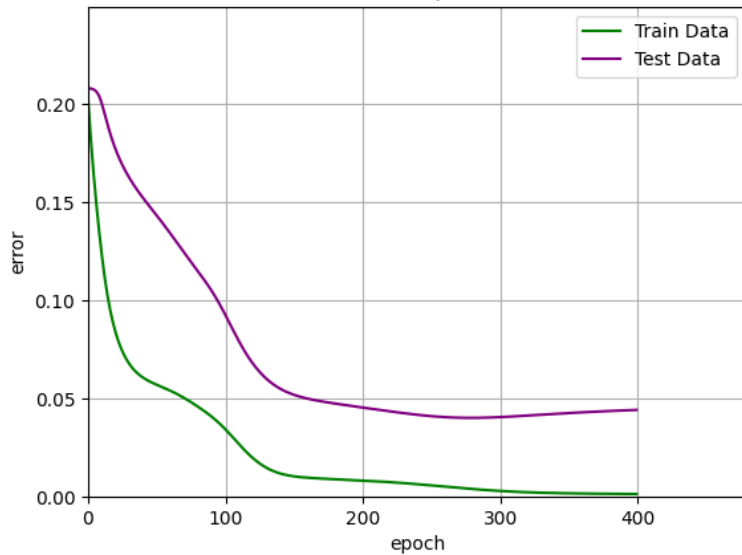
4) Οκτώ κέντρα, τυχαία επιλογή κέντρων, μικρό σ .



```
numHiddenLayerNeurons 8
numInputNeurons 53
numOutputNeurons 1
learningRates 0.5 0.1 0.4
sigmas 0.2
maxIterations 400
centresFile centers.txt
trainFile train.txt
testFile test.txt
```

399 6.945773436196886E-4 0.06558706786252001

Error - Epoch



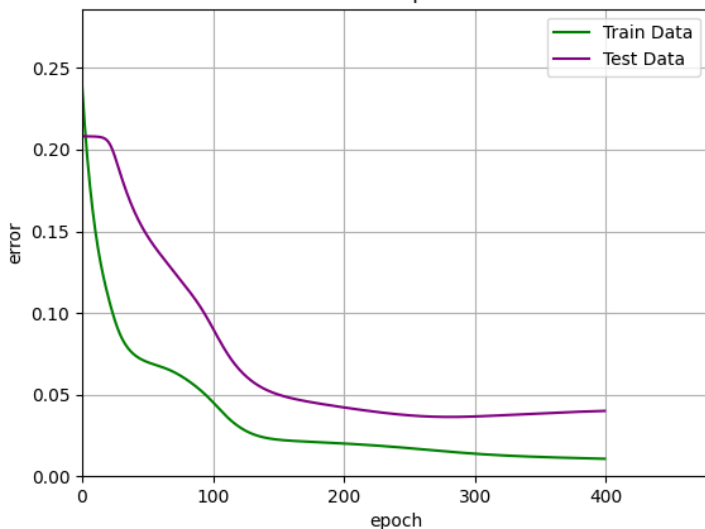
```
numHiddenLayerNeurons 8
numInputNeurons 53
numOutputNeurons 1
learningRates 0.05 0.01 0.04
sigmas 0.2
maxIterations 400
centresFile centers.txt
trainFile train.txt
testFile test.txt
```

399 0.0014742216541578717 0.044286343945117095

Παρατηρούμε ότι με σ ίσο με δύο, ναι μεν το σφάλμα εκπαίδευσης είναι εξαιρετικά χαμηλό, το δε σφάλμα ελέγχου δεν είναι ικανοποιητικά χαμηλό. Οπότε θα σταματήσουμε να δοκιμάζουμε με τόσο μεγάλο πλάτος. Βάση της θεωρίας οι επόμενες δοκιμές θα μας δώσουν τα βέλτιστα αποτελέσματα για κατά προσέγγιση λύση (λίγα κέντρα, «μεγάλο» σ , απλά χρειάζεται να εξερευνήσουμε το πόσο μεγάλο πρέπει να είναι το πλάτος). Από τα παραπάνω παραδείγματα καταλήγουμε στο συμπέρασμα ότι οι ρυθμοί μάθησης δουλεύουν καλύτερα αν έχουν μικρές τιμές – παράδειγμα 3.

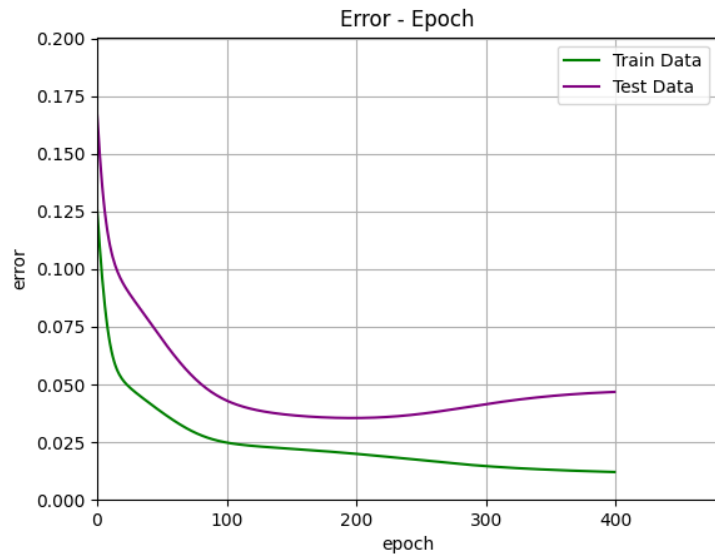
5) Τέσσερα κέντρα, τυχαία επιλογή κέντρων – Εξερεύνηση πλάτους γκαουσιανής.

Error - Epoch



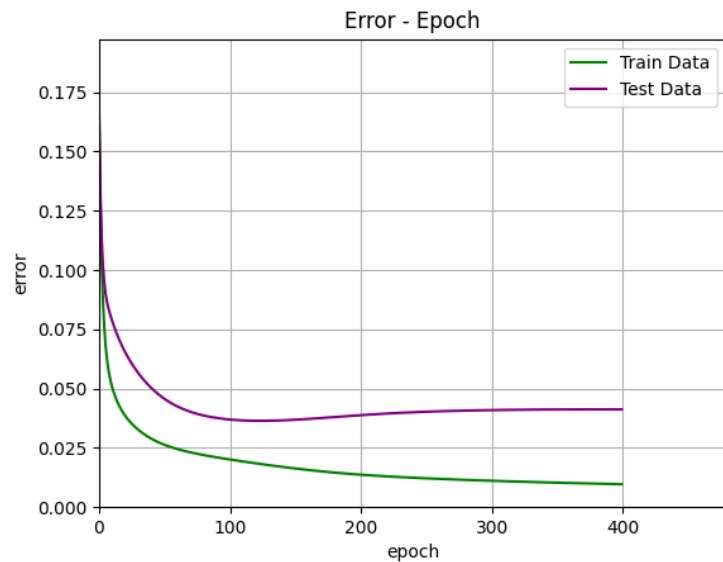
```
numHiddenLayerNeurons 4
numInputNeurons 53
numOutputNeurons 1
learningRates 0.05 0.01 0.04
sigmas 0.2
maxIterations 400
centresFile small.txt
trainFile train.txt
testFile test.txt
```

399 0.010686443151264642 0.04000658079265108



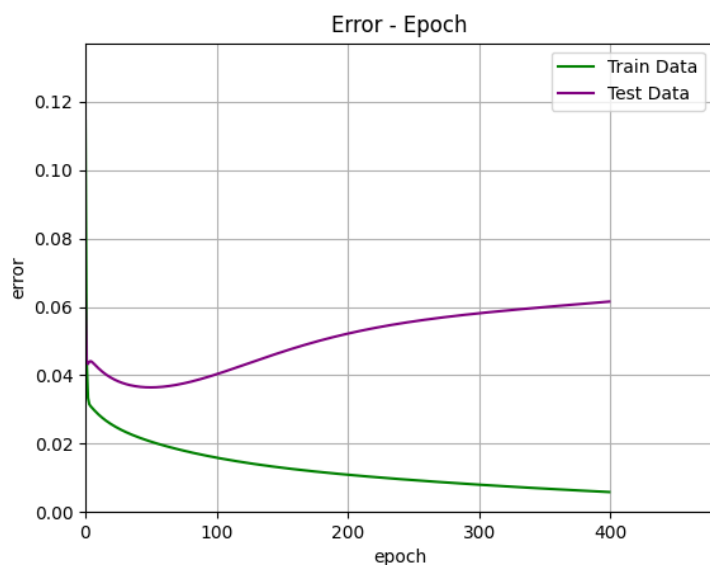
```
numHiddenLayerNeurons 4
numInputNeurons 53
numOutputNeurons 1
learningRates 0.05 0.01 0.04
sigmas 0.5
maxIterations 400
centresFile small.txt
trainFile train.txt
testFile test.txt
```

399 0.012077138364186253 0.04678167333186823



```
numHiddenLayerNeurons 4
numInputNeurons 53
numOutputNeurons 1
learningRates 0.05 0.01 0.04
sigmas 0.78
maxIterations 400
centresFile small.txt
trainFile train.txt
testFile test.txt
```

399 0.009661891340501473 0.04122011831333248



```
numHiddenLayerNeurons 4
numInputNeurons 53
numOutputNeurons 1
learningRates 0.05 0.01 0.04
sigmas 1.5
maxIterations 400
centresFile small.txt
trainFile train.txt
testFile test.txt
```

399 0.009661891340501473 0.04122011831333248

Ανάλυση Αποτελεσμάτων

Μετά από αρκετά τρεξίματα και δοκιμές του αλγορίθμου καταλήξαμε στα εξής συμπεράσματα.

- **Αριθμός επαναλήψεων:** Ο αριθμός επαναλήψεων δεν επηρεάζει σε πολύ μεγάλο βαθμό την επίδοση του δικτύου μας, εκτός βέβαια αν μιλάμε για πολύ μικρές ή πολύ μεγάλες τιμές όπου παρατηρούμε ανώμαλη συμπεριφορά.

Συμπέρασμα: Διατηρούμε τον αριθμό επαναλήψεων σε μια μέση τιμή (100 – 500).

- **Αριθμός κέντρων:** Αν κοιτάμε μόνο τις απόλυτες τιμές σφάλματος, στην πρώτη εκτέλεση βλέπουμε ότι έχουμε το μικρότερο σφάλμα (και εκπαίδευσης και ελέγχου). Ωστόσο βλέπουμε μεγάλη διαφορά μεταξύ εκπαίδευσης και ελέγχου, λογική συμπεριφορά βάσει της θεωρίας καταλήγουμε σε παρεμβολή επομένως δεν είναι ιδανικό για το σκοπό μας. Την ίδια συμπεριφορά παρατηρούμε και στα πειράματα με οκτώ κέντρα (μεγάλες αποκλίσεις). Στις περιπτώσεις με τα τέσσερα κέντρα βλέπουμε ότι ενώ το σφάλμα εκπαίδευσης δεν είναι τόσο χαμηλό όσο τις άλλες φορές, ωστόσο το σφάλμα ελέγχου είναι ελαφρώς πιο μικρό από τις άλλες περιπτώσεις και πιο κοντά στο σφάλμα ελέγχου.

Συμπέρασμα: Ανάλογα με τον στόχο που έχουμε επιλέγουμε διαφορετικό αριθμό κέντρων, στη δική μας περίπτωση θα προτιμήσουμε τα τέσσερα.

- **Ρυθμοί μάθησης:** Δεν μπορούμε να βγάλουμε να εξάγουμε κάποιο συγκεκριμένο συμπέρασμα από τα τρεξίματα μας. Μια παρατήρηση η οποία βέβαια δεν είναι καθολική είναι ότι με μεγάλους ρυθμούς μάθησης βλέπουμε τη δημιουργία κορυφών στη γραφική του σφάλματος.

Συμπέρασμα: Γενικά είναι πιο ασφαλής η χρήση μικρών ρυθμών μάθησης, ειδικά στο ρυθμό μάθησης του πλάτους της γκαουσιανής (μεγάλες τιμές δημιουργούν μεγάλες κορυφές).

- **Δεδομένα:** Στις αρχικές μας προσπάθειες δοκιμάσαμε να εκπαιδεύσουμε το δίκτυό μας με μη-επεξεργασμένα δεδομένα. Ωστόσο παρατηρήσαμε ότι το δίκτυο δεν μπορούσε να γενικεύσει τη γνώση του στα δεδομένα ελέγχου και ειδικά στις αρνητικές τιμές βιολογικής δράσης.

Συμπέρασμα: Όπως και στις περισσότερες περιπτώσεις, είναι καλό και ίσως αναγκαίο να κανονικοποιούμε τα δεδομένα μας για να μην αφήνουμε ορισμένες τιμές να αποκτούν τόσο μεγάλη επιρροή στη μάθηση του δικτύου μας.

Training & Testing sets

Τα αρχεία εκπαίδευσης και ελέγχου δημιουργήθηκαν τυχαία, ακολουθώντας τη τακτική 70-30.