

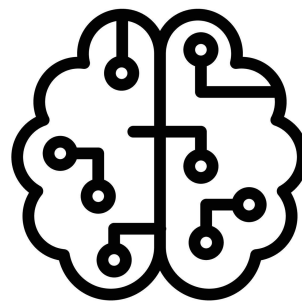
Εργασία 4 στα Ασαφή συστήματα

# Υπολογιστική Νοημοσύνη

Εκπαίδευση Μοντέλων TSK - Classification

Παναγιώτης Αβραμίδης 8768

Οκτώβριος 2024



Fuzzy Logic

Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης  
Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών  
Ηλεκτρονικών Υπολογιστών  
Τομέας Ηλεκτρονικής

# Περιεχόμενα

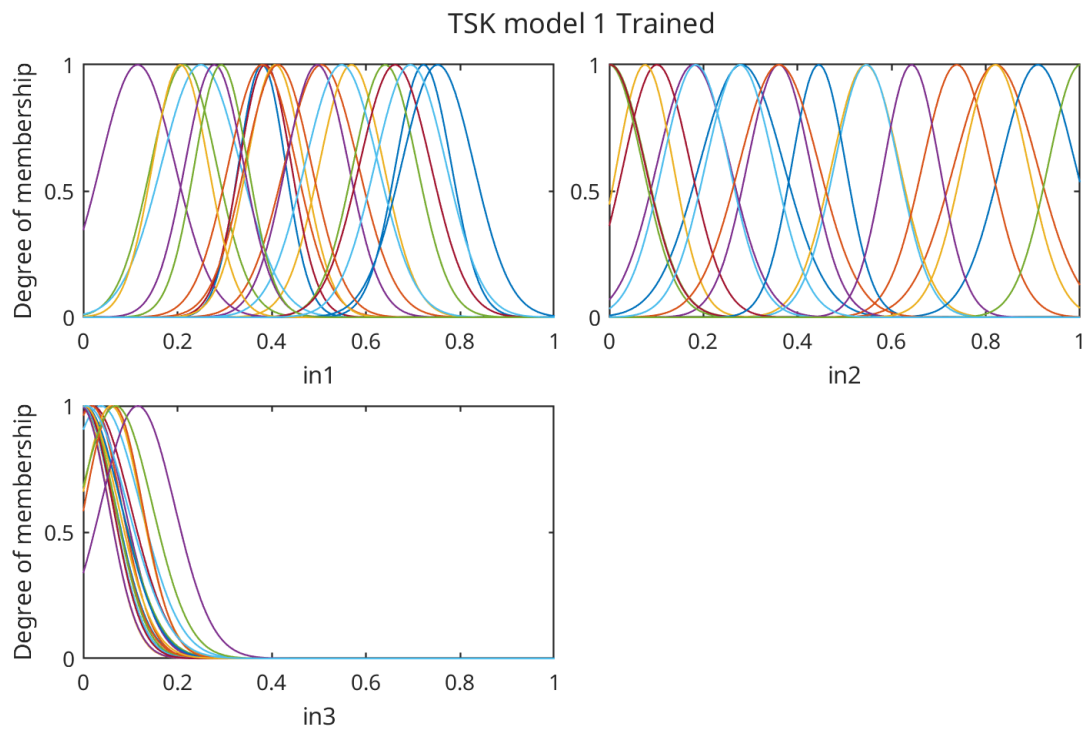
|          |                                  |           |
|----------|----------------------------------|-----------|
| <b>1</b> | <b>Περιγραφή Εργασίας</b>        | <b>1</b>  |
| <b>2</b> | <b>Εκπαίδευση 4 TSK Μοντέλων</b> | <b>2</b>  |
| 2.1      | TSK Μοντέλο 1 . . . . .          | 2         |
| 2.2      | TSK Μοντέλο 2 . . . . .          | 6         |
| 2.3      | TSK Μοντέλο 3 . . . . .          | 11        |
| 2.4      | TSK Μοντέλο 4 . . . . .          | 16        |
| 2.5      | Σχολιασμός . . . . .             | 21        |
| <b>3</b> | <b>Μοντέλο Υψηλών Διαστάσεων</b> | <b>22</b> |
| 3.1      | K-fold . . . . .                 | 22        |
| 3.2      | Αποτελέσματα . . . . .           | 23        |
|          | <b>Βιβλιογραφία</b>              | <b>28</b> |

# 1 Περιγραφή Εργασίας

Ζητείται η υλοποίηση μοντέλων TSK (Takagi-Sugeno-Kang). Στο πρώτο μέρος θα πρέπει το κάθε μοντέλο να διαμεριστεί με τη μέθοδο Subtractive Clustering όπου το clustering θα γίνει είτε σε όλο το dataset είτε ανά κλάση και με δύο ακραίες τιμές για την ακτίνα των clusters κάθε φορά (Πίνακας). Στο δεύτερο μέρος επειδή το dataset έχει μεγάλες διαστάσεις, υπάρχει το φαινόμενο του Rule Explosion που είναι γενικότερο αποτέλεσμα της κατάρας της διαστασιμότητας (Curse of Dimensionality). Ως εκ τούτου εφαρμόζεται Subtractive Clustering για κάθε κλάση. Για τον υπολογισμό κλάσης, επειδή η κάθε κλάση εκφράζεται από διακριτούς αριθμούς χρησιμοποιείται διακριτοποίηση της εξόδου.

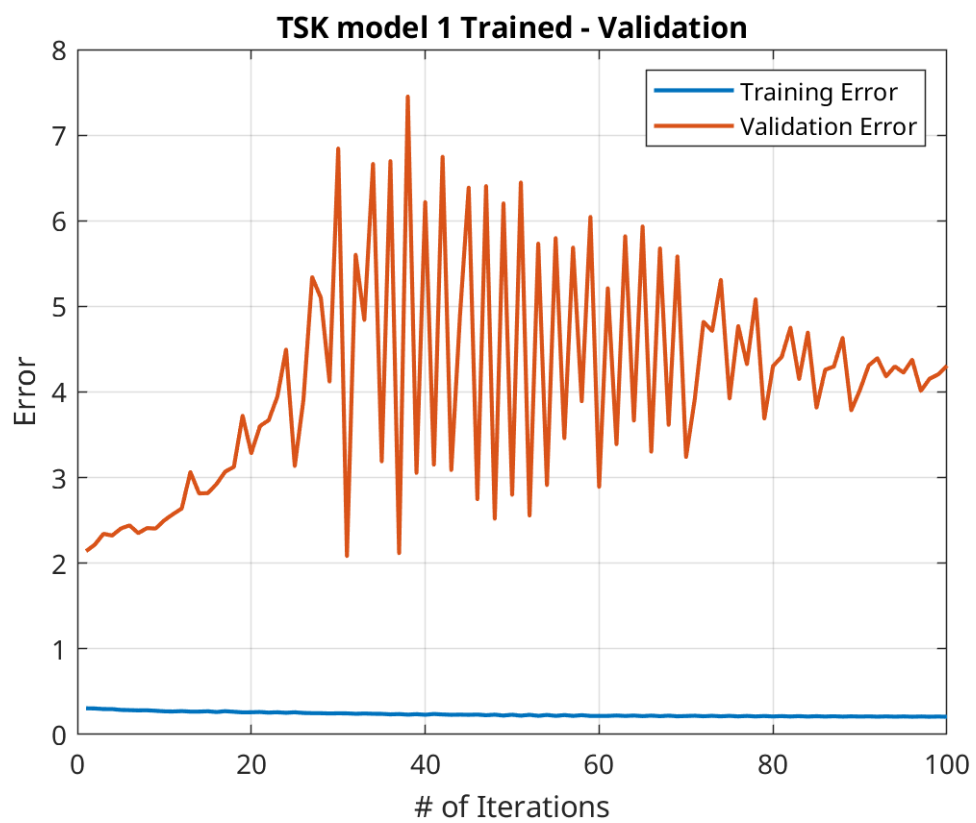


Μετά την εκπαίδευση οι συναρτήσεις συμμετοχής διαμορφώνονται ως εξής:

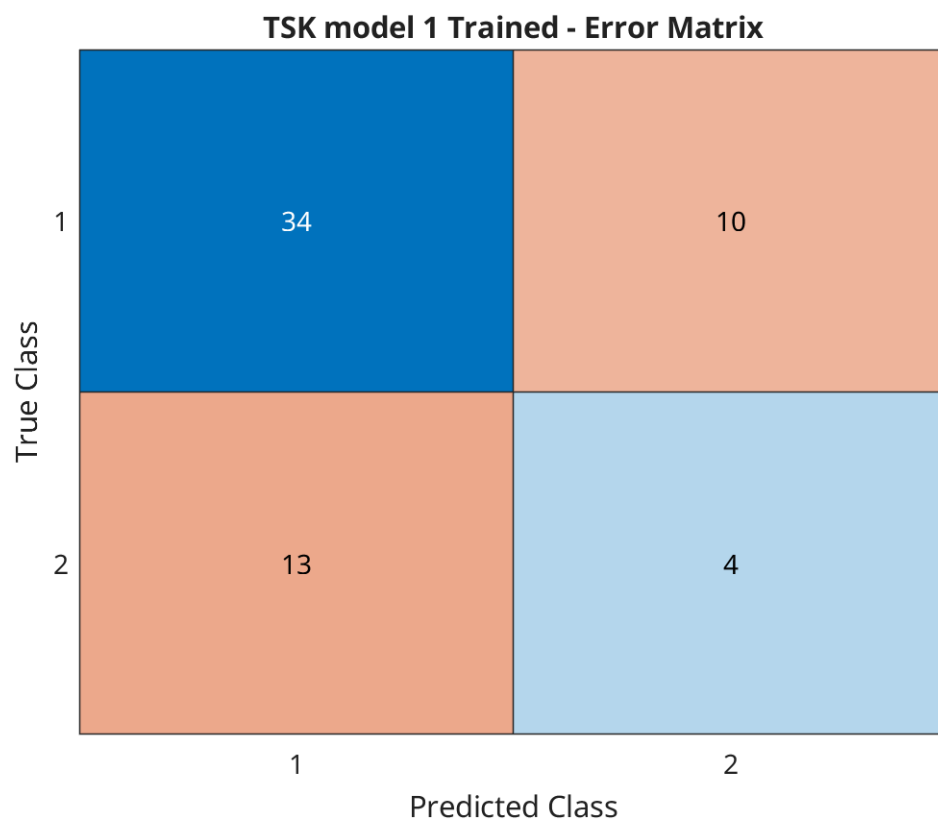


**Σχήμα 2.3:** Συναρτήσεις Συμμετοχής TSK 1 (Trained)

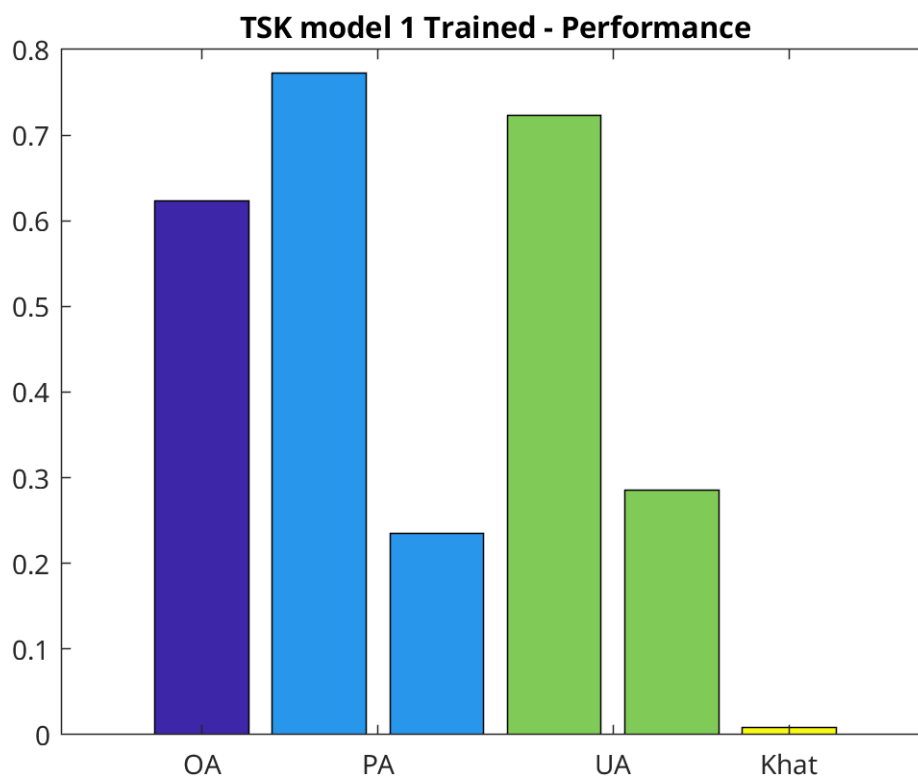
Η καμπύλη μάθησης του μοντέλου κατά την εκπαίδευση φαίνεται ως εξής:



**Σχήμα 2.4:** Σφάλμα Εκπαίδευσης και Validation Set (TSK 1)



Σχήμα 2.5: Ερρορ Ματριξ - TSK 1

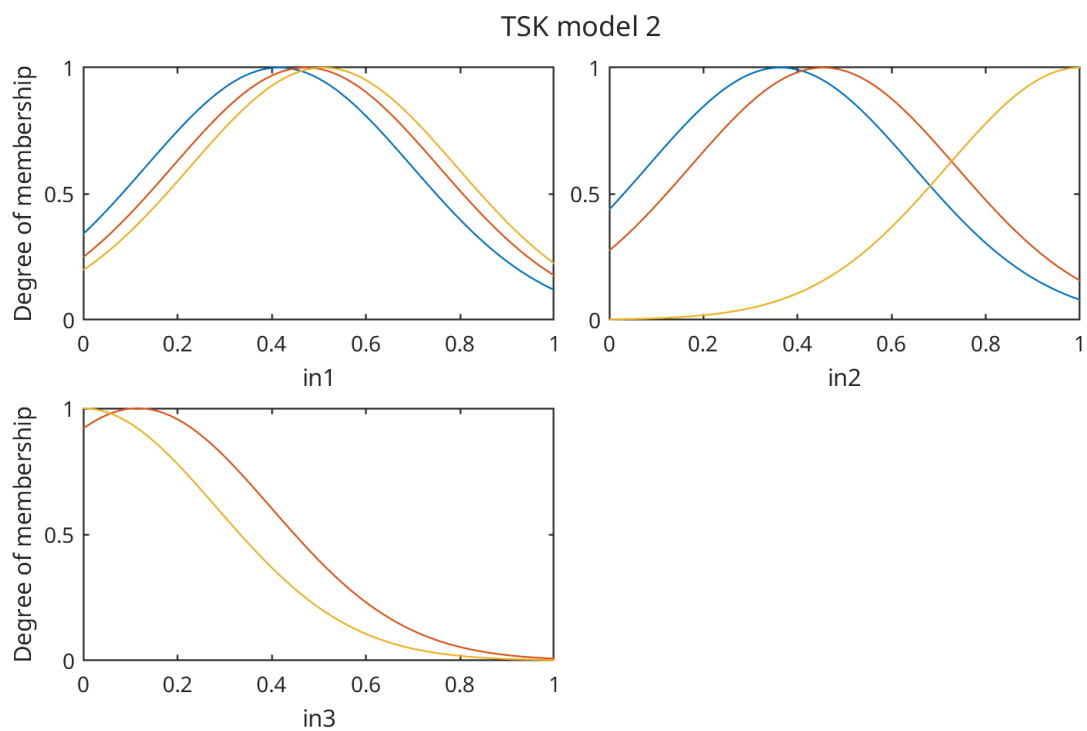


Σχήμα 2.6: Μετρικά απόδοσης - TSK 1

## 2.2 TSK Μοντέλο 2

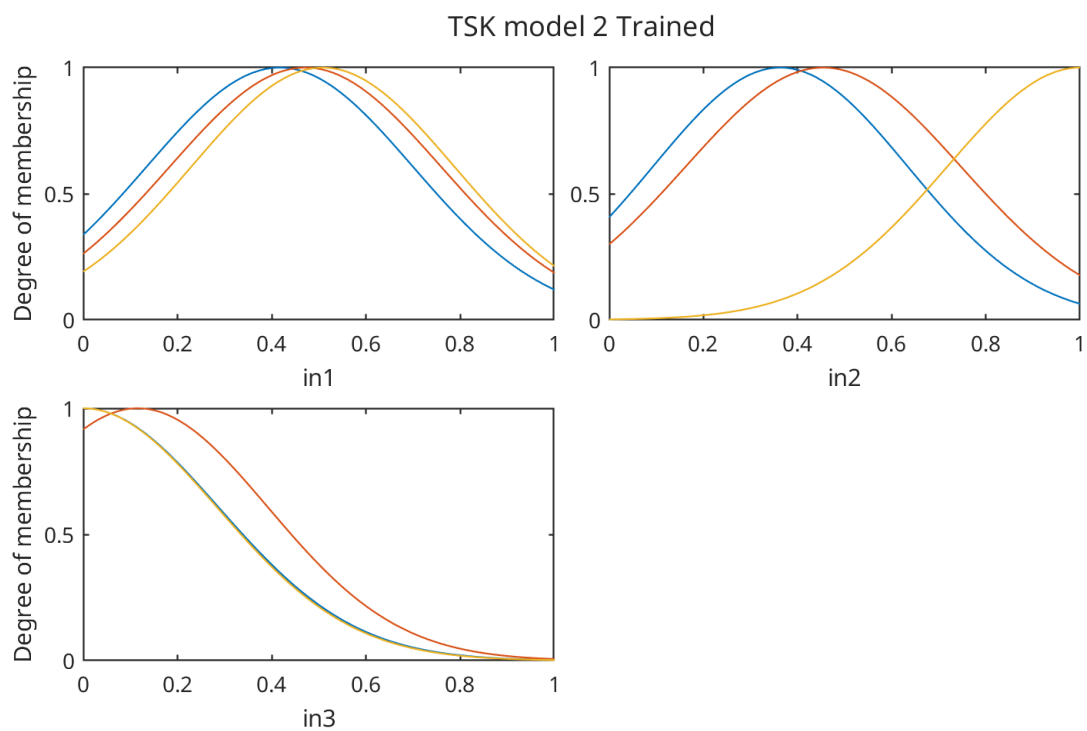
Για το δεύτερο μοντέλο τα αποτελέσματα είναι τα εξής:





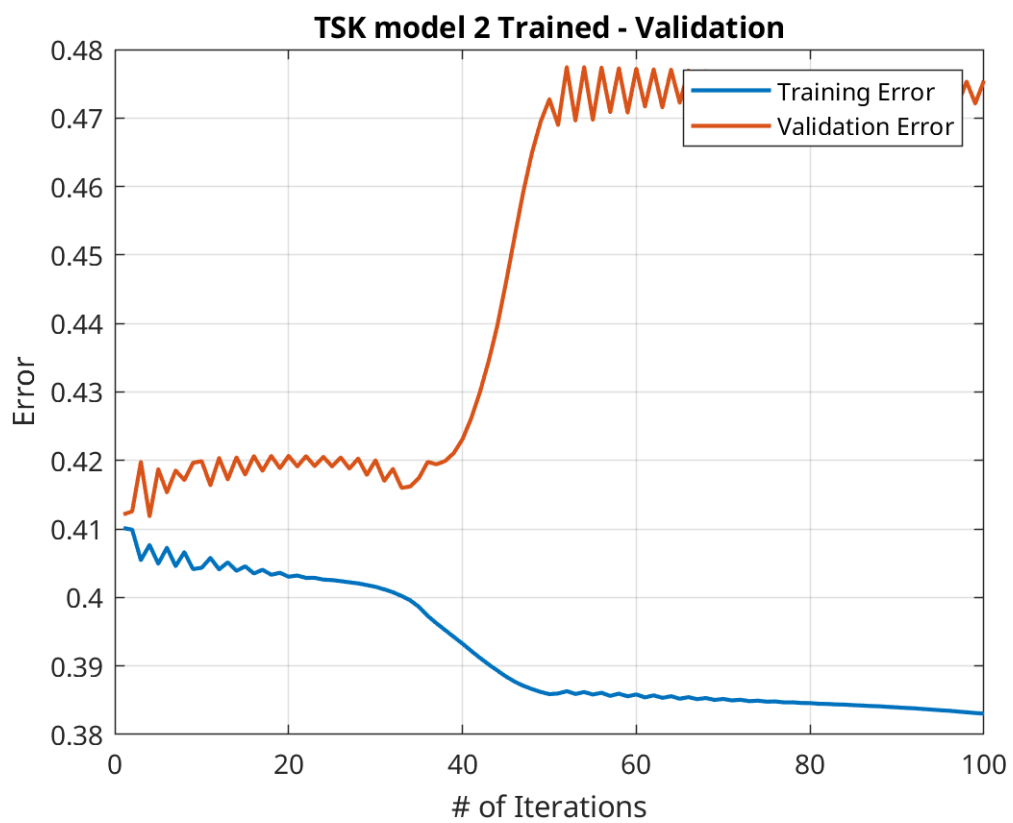
**Σχήμα 2.7:** Συναρτήσεις Συμμετοχής TSK 2

Μετά την εκπαίδευση οι συναρτήσεις συμμετοχής διαμορφώνονται ως εξής:

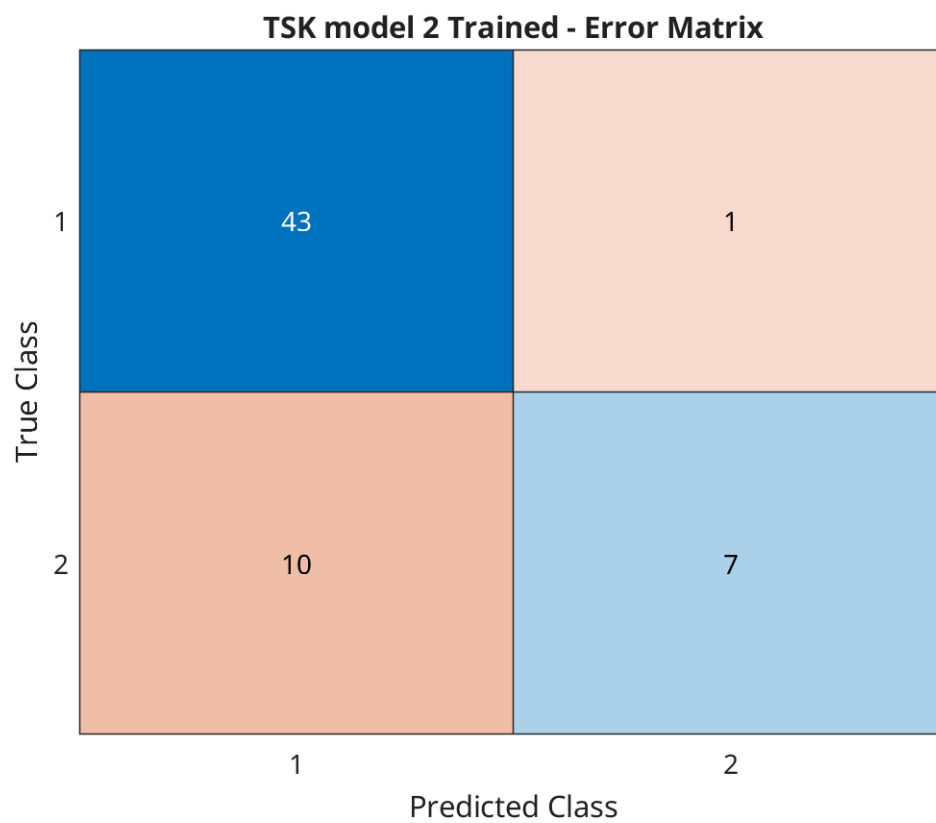


**Σχήμα 2.8:** Συναρτήσεις Συμμετοχής TSK 2 (Trained)

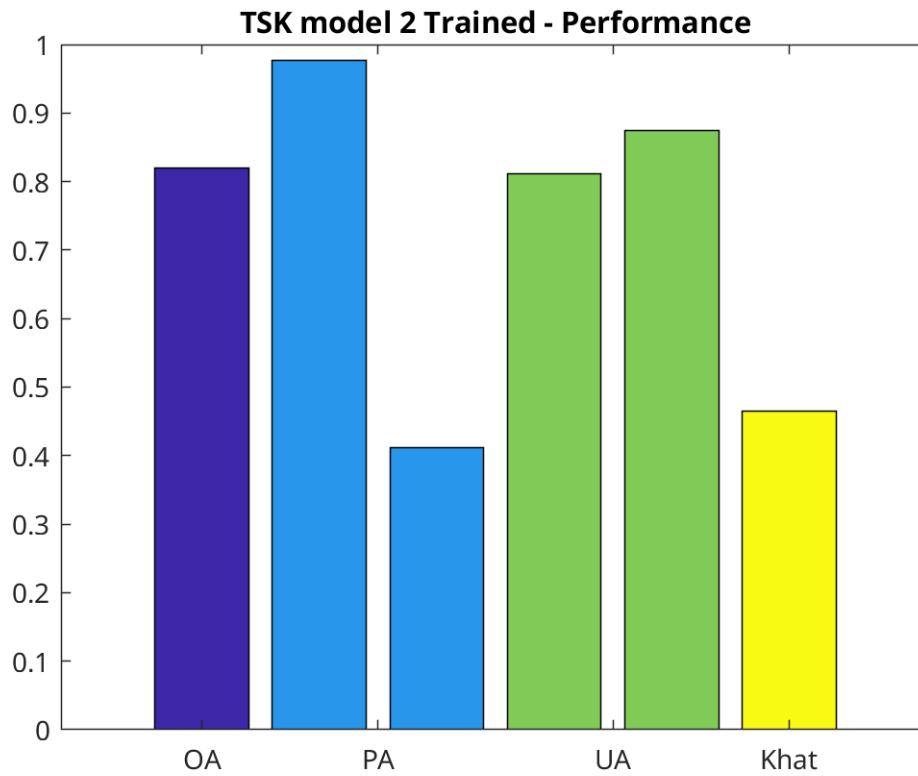
Η καμπύλη μάθησης του μοντέλου κατά την εκπαίδευση φαίνεται ως εξής:



Σχήμα 2.9: Σφάλμα Εκπαίδευσης και Validation Set (TSK 2)



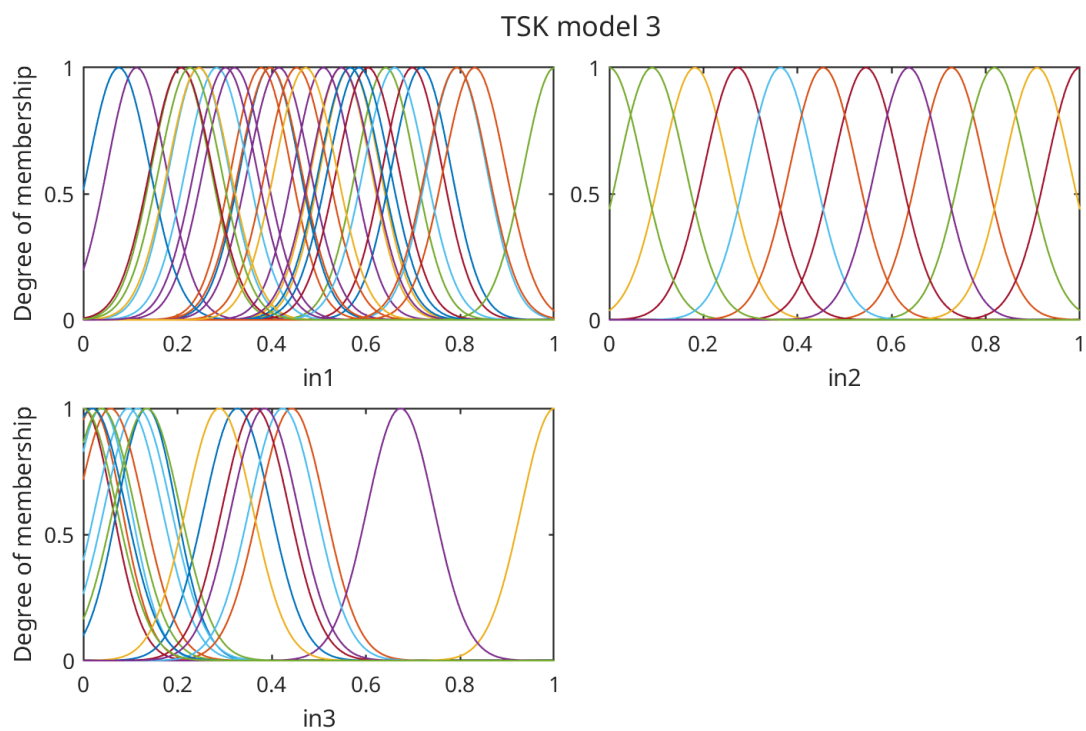
**Σχήμα 2.10:** Ερρορ Ματριζ - TSK 2



Σχήμα 2.11: Μετρικά απόδοσης - TSK 2

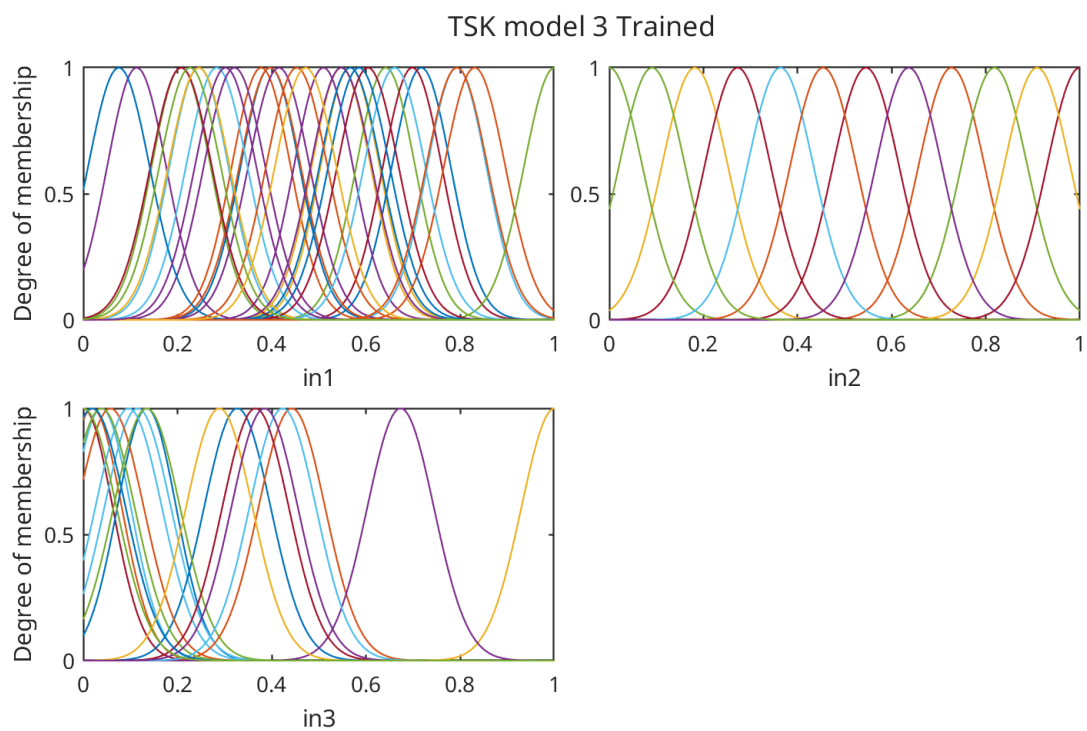
## 2.3 TSK Μοντέλο 3

Για το τρίτο μοντέλο τα αποτελέσματα είναι τα εξής:



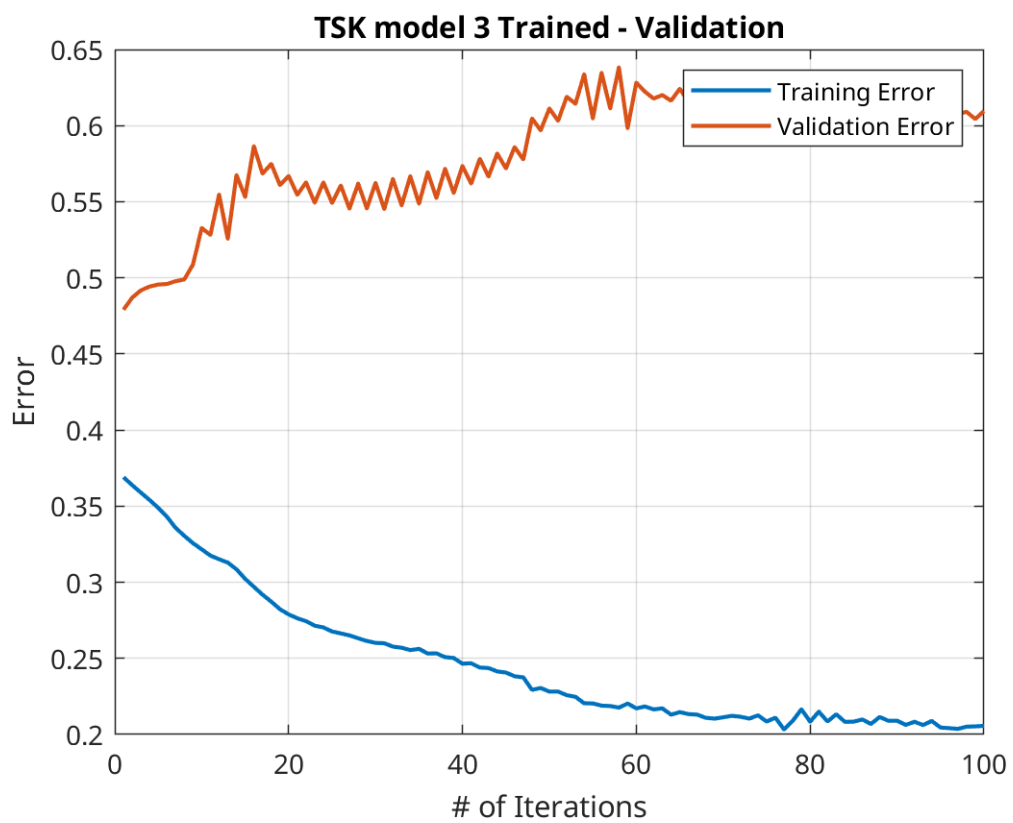
**Σχήμα 2.12:** Συναρτήσεις Συμμετοχής TSK 3

Μετά την εκπαίδευση οι συναρτήσεις συμμετοχής διαμορφώνονται ως εξής:



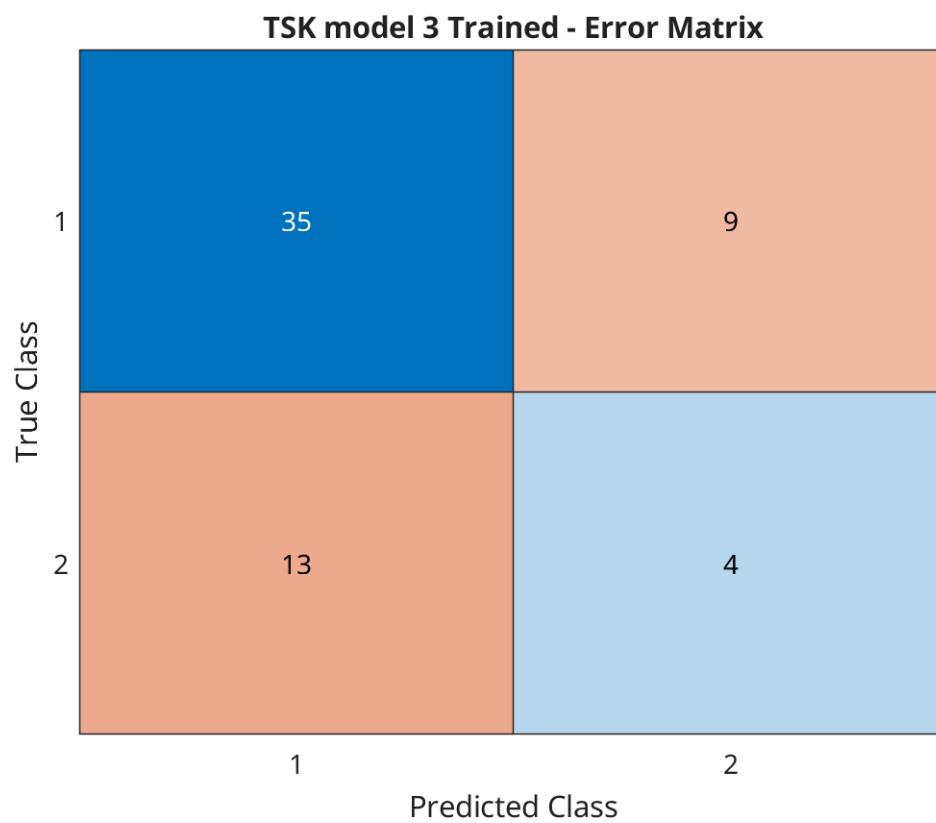
**Σχήμα 2.13:** Συναρτήσεις Συμμετοχής TSK 3 (Trained)

Η καμπύλη μάθησης του μοντέλου κατά την εκπαίδευση φαίνεται ως εξής:

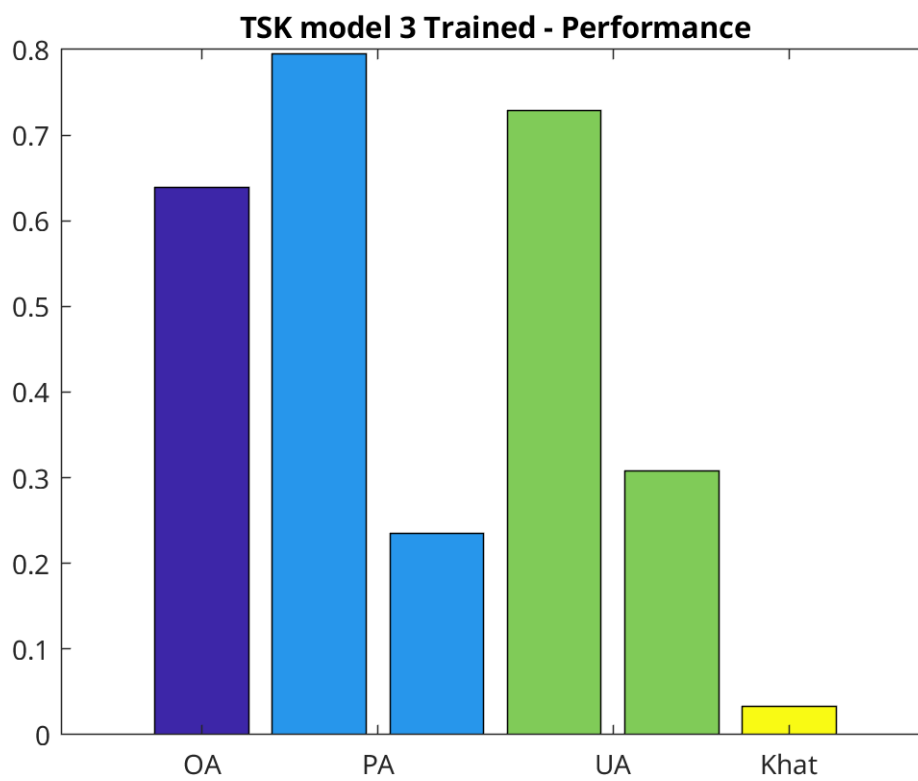


Σχήμα 2.14: Σφάλμα Εκπαίδευσης και Validation Set (TSK 3)





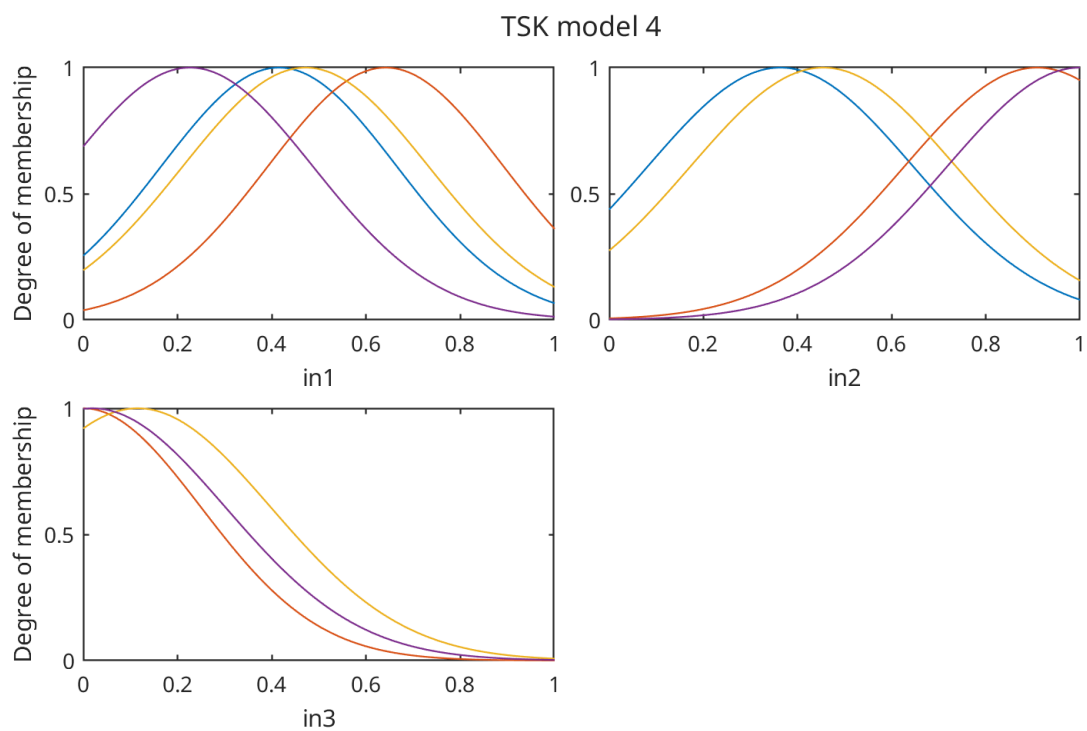
Σχήμα 2.15: Ερρορ Ματριζ - TSK 3



Σχήμα 2.16: Μετρικά απόδοσης - TSK 3

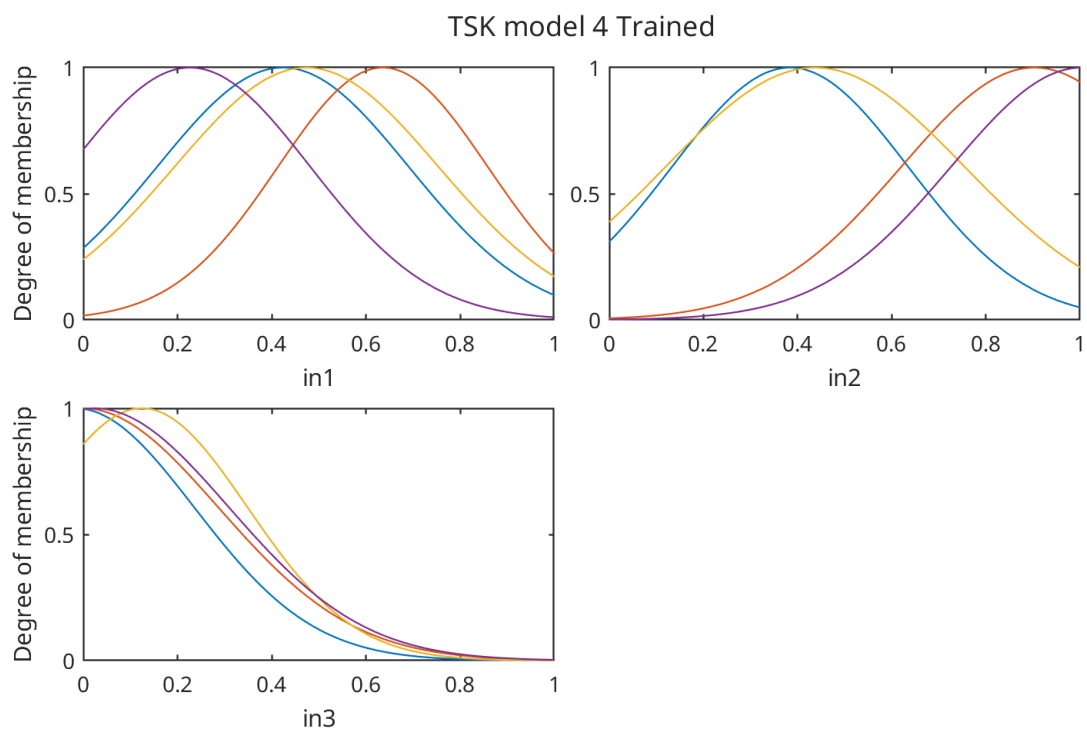
## 2.4 TSK Μοντέλο 4

Για το τέταρτο μοντέλο τα αποτελέσματα είναι τα εξής:



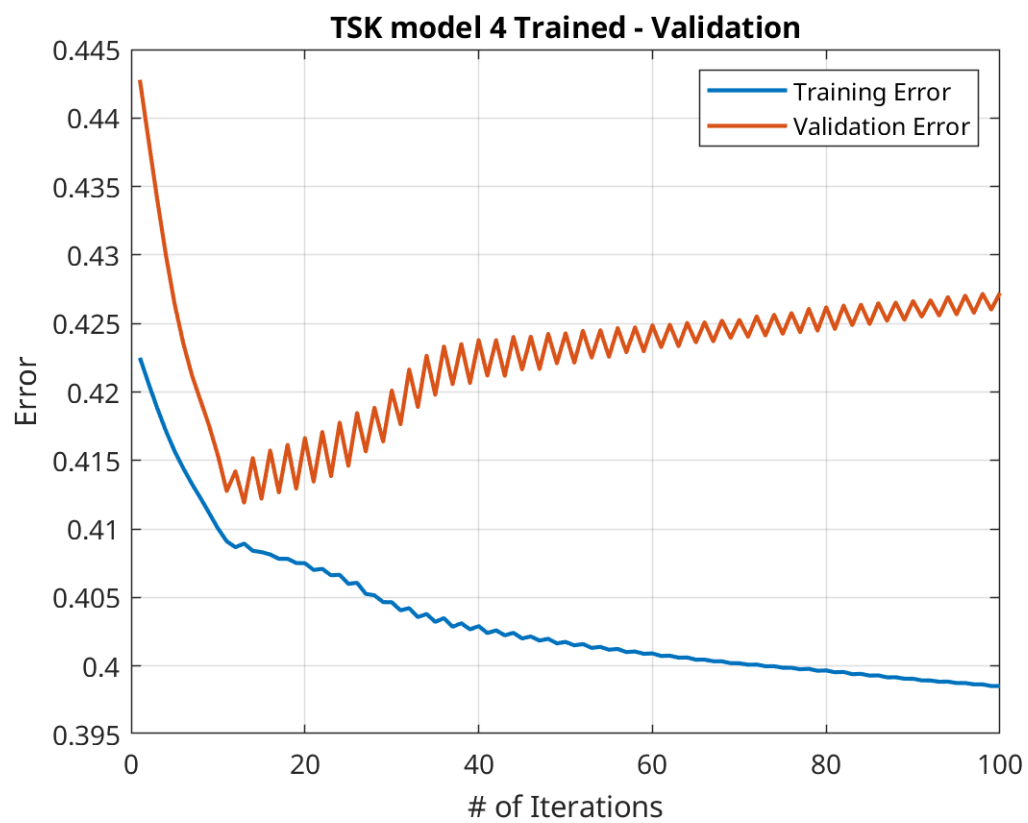
**Σχήμα 2.17:** Συναρτήσεις Συμμετοχής TSK 4

Μετά την εκπαίδευση οι συναρτήσεις συμμετοχής διαμορφώνονται ως εξής:

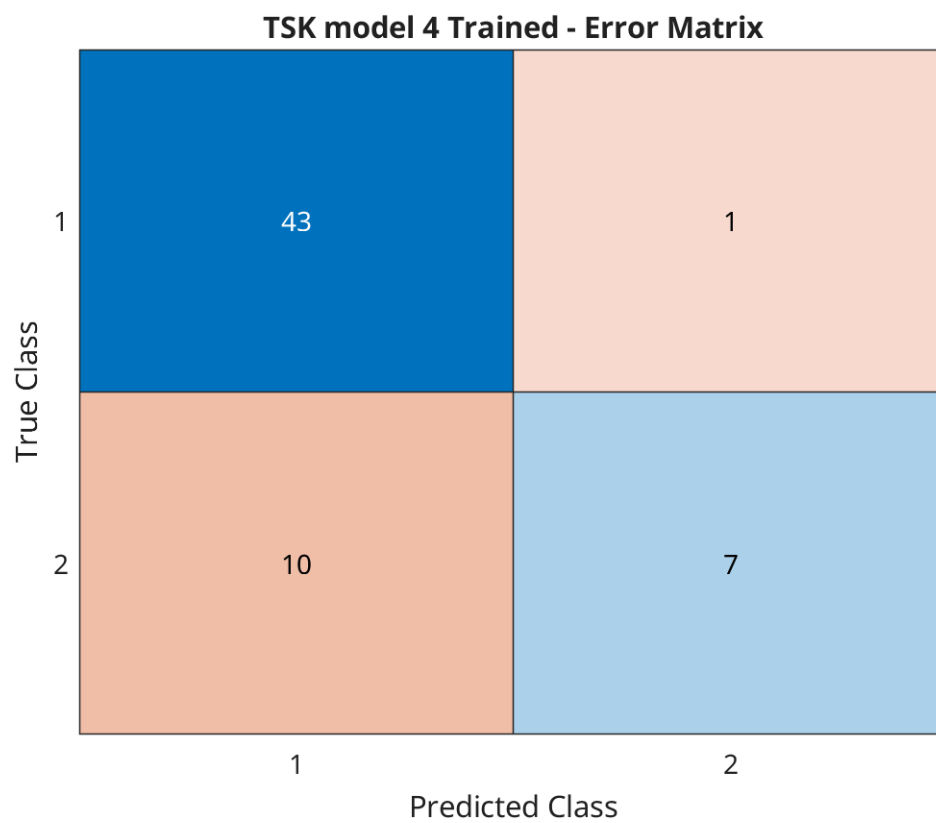


**Σχήμα 2.18:** Συναρτήσεις Συμμετοχής TSK 4 (Trained)

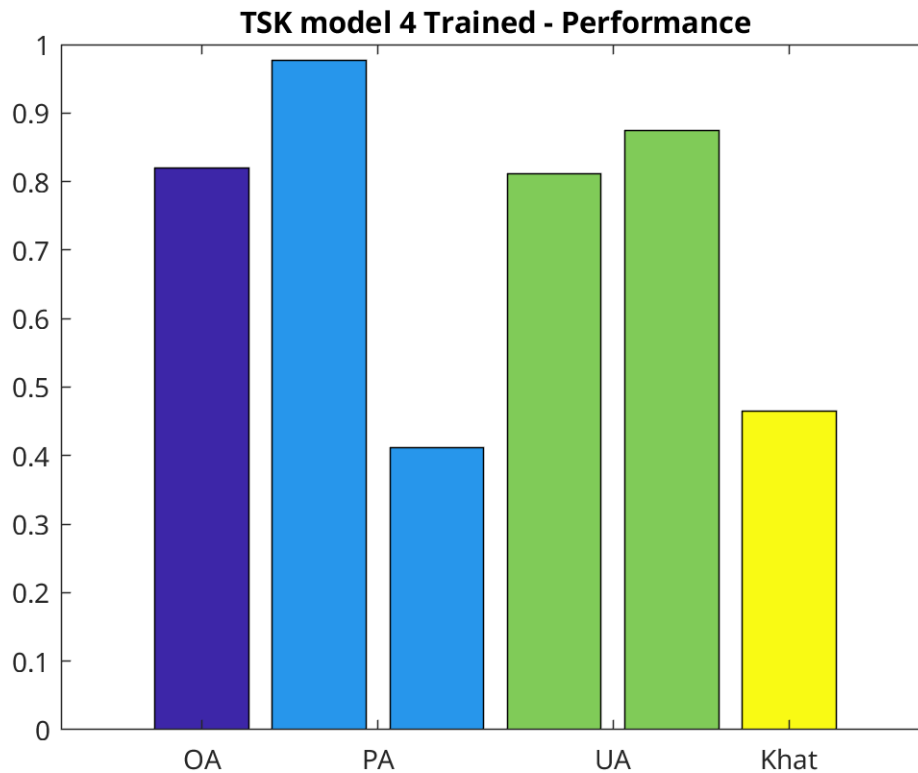
Η καμπύλη μάθησης του μοντέλου κατά την εκπαίδευση φαίνεται ως εξής:



Σχήμα 2.19: Σφάλμα Εκπαίδευσης και Validation Set (TSK 4)



**Σχήμα 2.20:** Ερρορ Ματριζ - TSK 4



Σχήμα 2.21: Μετρικά απόδοσης - TSK 4

## 2.5 Σχολιασμός

Όπως φαίνεται από τον παρακάτω πίνακα το 4ο μοντέλο έχει το καλύτερο overall accuracy. Είναι το Class Dependent μοντέλο με τη μεγάλη ακτίνα.

| OA      | PA      | UA      | Khat      |
|---------|---------|---------|-----------|
| 0.62295 | 0.50401 | 0.50456 | 0.0084806 |
| 0.81967 | 0.69452 | 0.84316 | 0.46449   |
| 0.63934 | 0.51537 | 0.51843 | 0.033141  |
| 0.81967 | 0.69452 | 0.84316 | 0.46449   |

Σχήμα 2.22: Πίνακας μετρικών των τεσσάρων μοντέλων

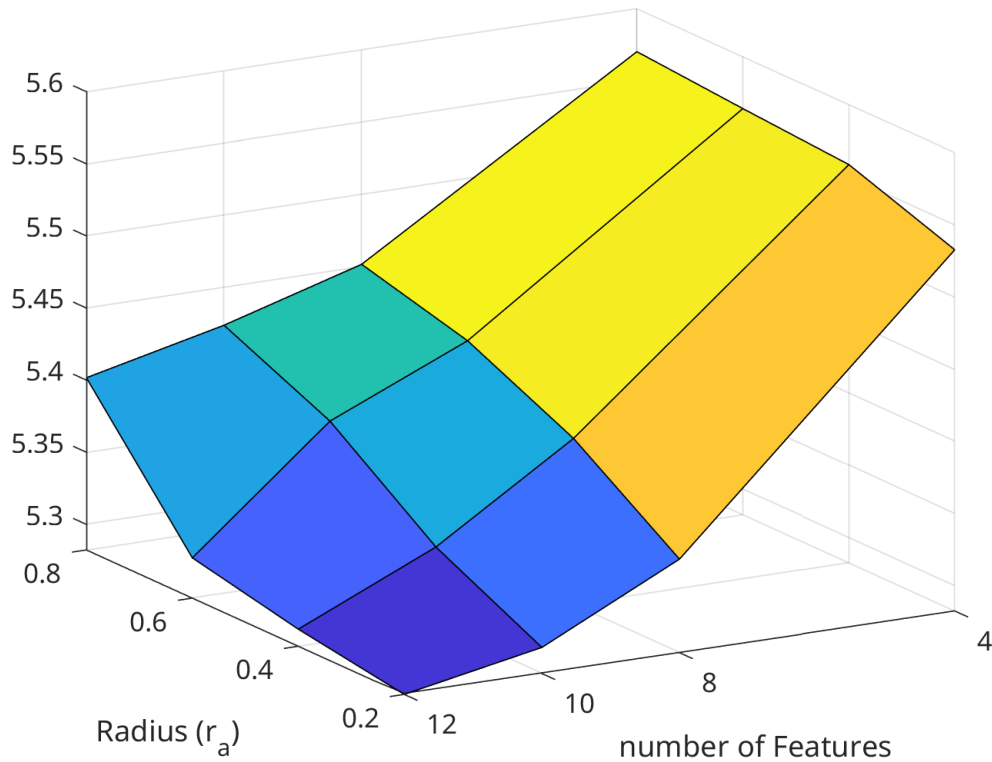
## 3 Μοντέλο Υψηλών Διαστάσεων

Το δεύτερο μέρος της εργασίας ζητάει να εκπαιδευτεί ένα μοντέλο με 178 διαστάσεις (χαρακτηριστικά), και για αυτό να χρησιμοποιηθεί Subtractive Clustering και μείωση διαστάσεων. Για τη μείωση διαστάσεων χρησιμοποιήθηκε ο αλγόριθμος Relief, διότι ενώ είναι πιο χρονοβόρος, αποδίδει καλύτερα από τον mRMR. Προκειμένου να επιλεγούν οι υπερπαράμετροι της μεθόδου αυτής (αριθμός features) και ακτίνα των cluster εφαρμόζεται Grid Search 5-φολδ με αξιολόγηση τον μέσο όρο του σφάλματος του Validation Error ώστε να έχω μια αντικειμενική εικόνα για το πόσο καλή επιλογή είναι οι εκάστοτε παράμετροι. Το grid search έγινε για ακτίνες [0.2 0.4 0.6 0.8 ] και αριθμό χαρακτηριστικών [4 8 10 12]. Ο λόγος που επιλέχθηκαν τουλάχιστον 4 χαρακτηριστικά είναι γιατί όπως και στην προηγούμενη εργασία παρατηρήθηκε ότι η απόδοση του μοντέλου αυξάνεται καθώς αυξάνονται τα χαρακτηριστικά, που είναι και λογικό αφού το μοντέλο μπορεί να μάθει από περισσότερη πληροφορία.

### 3.1 K-fold

Η αναζήτηση Grid Search (5-φολδ) έχει το εξής σφάλμα ως προς τον αριθμό χαρακτηριστικών και την ακτίνα των cluster. Από το γράφημα φαίνεται ότι το σφάλμα φθίνει καθώς φθίνει το  $r_a$  ή αυξάνει ο αριθμός χαρακτηριστικών. Τελικά το προς εκπαίδευση μοντέλο είναι αυτό με  $r_a = 0.2$  και num\_features=12.

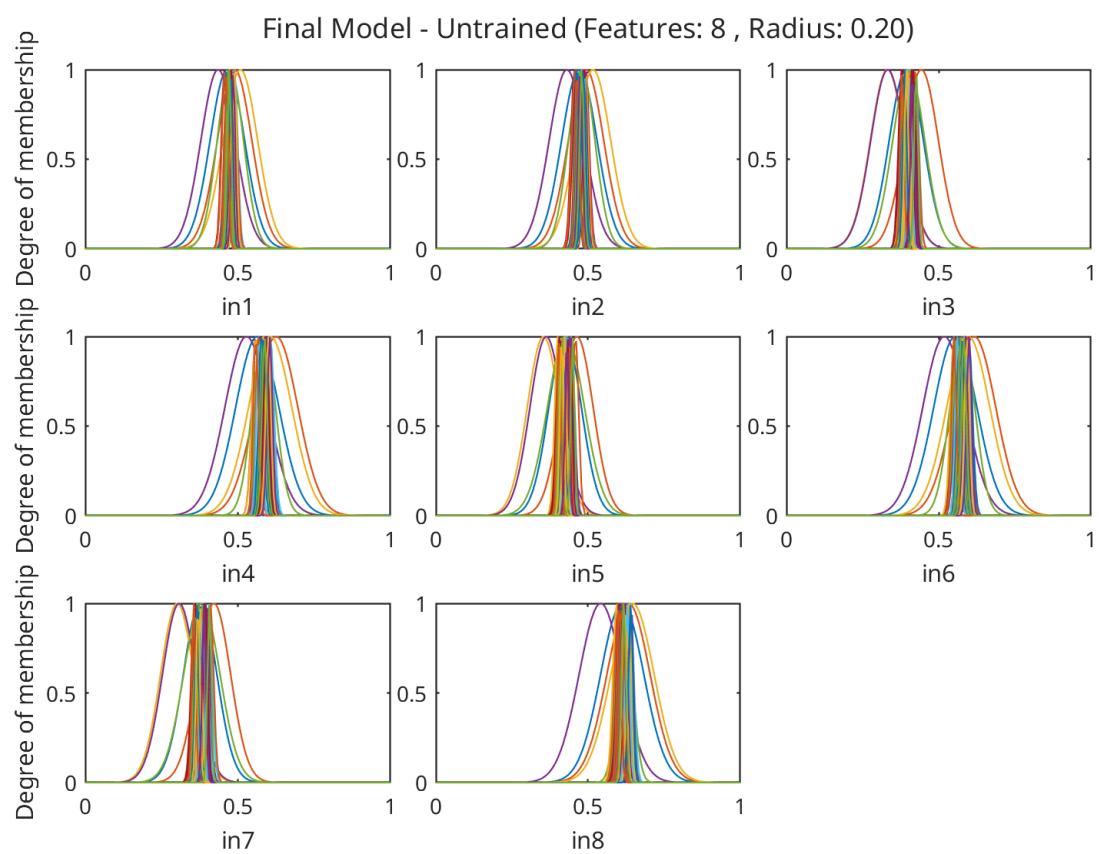




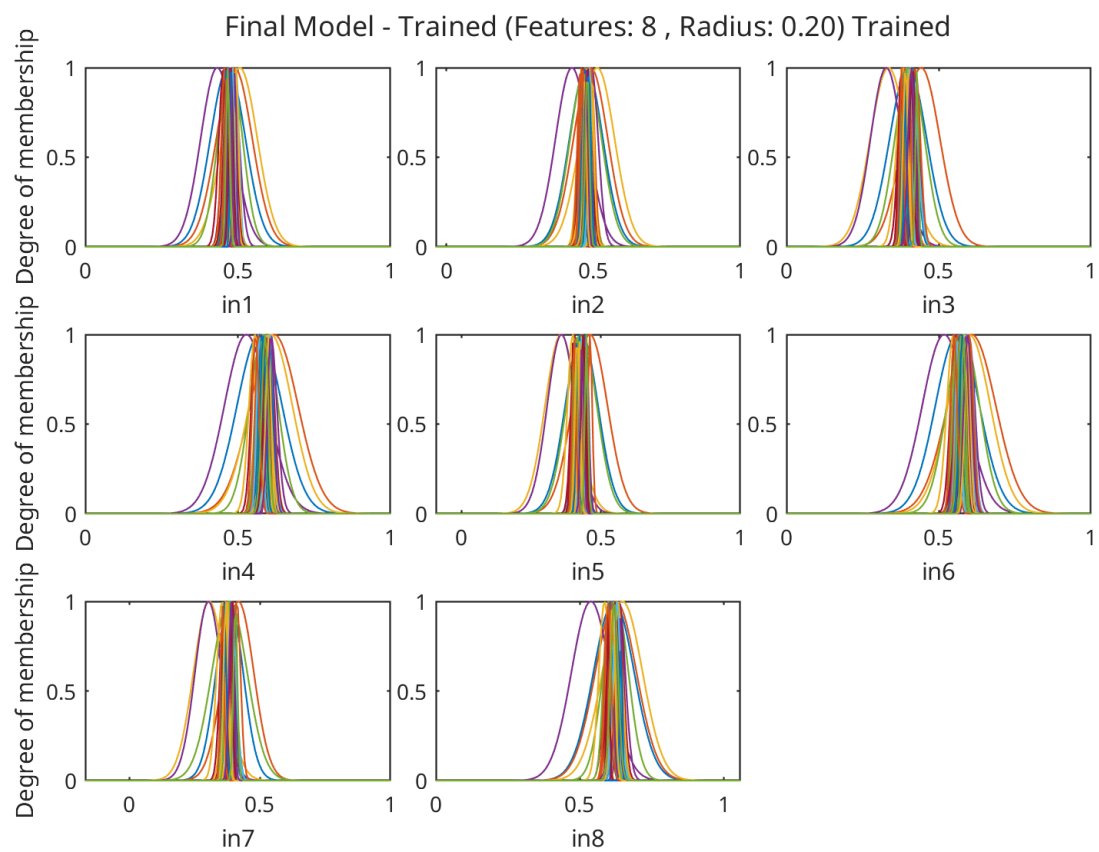
Σχήμα 3.1: Μέσο Validation Error

## 3.2 Αποτελέσματα

Ακολουθούν οι συναρτήσεις συμμετοχής πριν και μετά την εκπαίδευση. Όπως φαίνεται, το Overall Accuracy φτάνει το 0.4. Αν και πήρε αρκετή ώρα το Grid Search (σχεδόν 2 ώρες) τελικά κρίνεται ότι συμφέρει γιατί ο αριθμός των κανόνων πέφτει δραματικά, καθώς αποφεύγεται το Rule Explosion. Ενδεχόμενη εκπαίδευση με όλα τα χαρακτηριστικά θα είχε ως αποτέλεσμα την εκτίναξη του χρόνου εκπαίδευσης σε μη λογικό χρόνο.



**Σχήμα 3.2:** Συναρτήσεις Συμμετοχής (αρχικά)



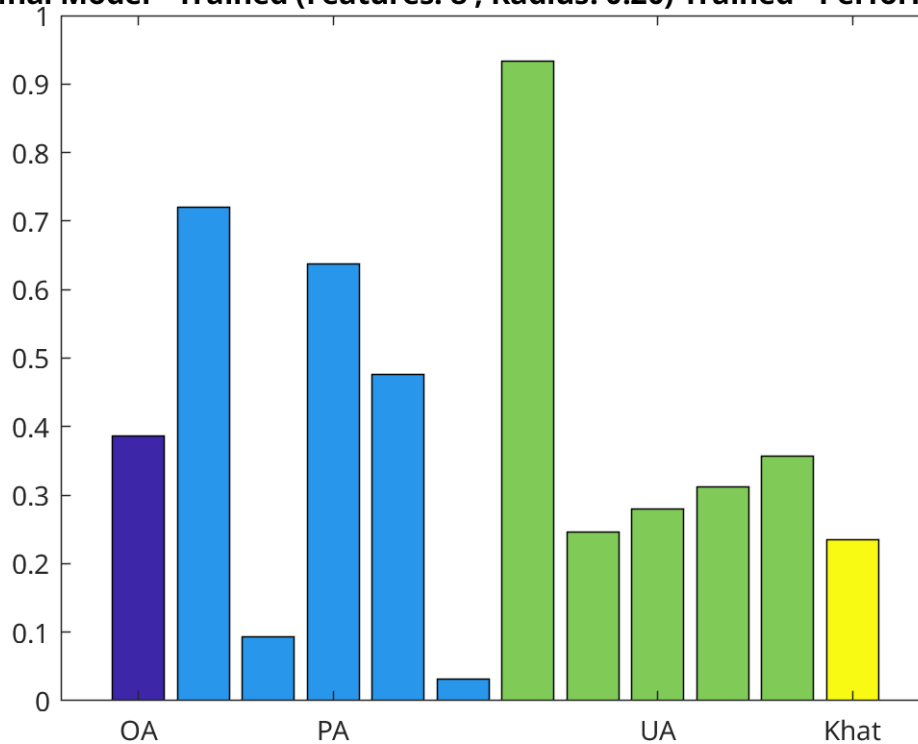
**Σχήμα 3.3:** Συναρτήσεις Συμμετοχής (τελικά)

**Final Model - Trained (Features: 8 , Radius: 0.20) Trained - Error Matrix**

|            |   |                 |    |     |     |    |
|------------|---|-----------------|----|-----|-----|----|
| True Class | 1 | 322             | 70 | 41  | 13  | 1  |
|            | 2 | 20              | 43 | 289 | 102 | 4  |
|            | 3 | 1               | 25 | 285 | 125 | 11 |
|            | 4 | 2               | 32 | 202 | 224 | 11 |
|            | 5 |                 | 5  | 202 | 255 | 15 |
|            |   | 1               | 2  | 3   | 4   | 5  |
|            |   | Predicted Class |    |     |     |    |

**Σχήμα 3.4:** Σφάλματα Πρόβλεψης - (swarmplot)

**Final Model - Trained (Features: 8 , Radius: 0.20) Trained - Performance**



**Σχήμα 3.5:** Σφάλματα Πρόβλεψης - (swarmplot)

| OA     | PA     | UA     | Khat   |
|--------|--------|--------|--------|
| 0.4065 | 0.4112 | 0.4769 | 0.2594 |

**Σχήμα 3.6:** Πίνακας μετρικών του τελικού μοντέλου

# Βιβλιογραφία

- [1] Ασαφή Συστήματα, Ι. Θεοχάρης <https://elearning.auth.gr/mod/folder/view.php?id=446790>