

Υπολογιστική Νοημοσύνη: TSK Regression

Τσακιδίδου Δήμητρα Μαρία - 10466

Σεπτέμβριος 2025

Μέρος 1: Εφαρμογή σε μικρό dataset

Περιγραφή Δεδομένων και Προεπεξεργασία

Για το πρώτο μέρος της εργασίας χρησιμοποιήθηκε το σύνολο δεδομένων `airfoil_self_noise.dat` από το UCI repository, το οποίο περιλαμβάνει 1503 δείγματα με 6 χαρακτηριστικά (5 εισόδους και 1 έξοδο).

Τα δεδομένα διαχωρίστηκαν σε τρία υποσύνολα:

- Σύνολο εκπαίδευσης (60% των δεδομένων)
- Σύνολο επικύρωσης (20% των δεδομένων)
- Σύνολο ελέγχου (20% των δεδομένων)

Εφαρμόστηκε κανονικοποίηση των δεδομένων με τη συνάρτηση `split.scale` για βελτίωση της απόδοσης των μοντέλων.

Μοντέλα TSK

Εκπαιδεύτηκαν τέσσερα μοντέλα TSK με διαφορετικές παραμέτρους:

Table 1: Παράμετροι TSK μοντέλων

Μοντέλο	Πλήθος Συναρτήσεων Συμμετοχής	Μορφή Εξόδου
TSK 1	2	Singleton
TSK 2	3	Singleton
TSK 3	2	Polynomial
TSK 4	3	Polynomial

Αποτελέσματα

Τα αποτελέσματα των μοντέλων στο σύνολο ελέγχου παρουσιάζονται στον παρακάτω πίνακα:

Table 2: Μετρικές απόδοσης μοντέλων

Μοντέλο	RMSE	NMSE	NDEI	R^2
TSK 1	3.8397	0.3379	0.5813	0.6621
TSK 2	3.5006	0.2808	0.5299	0.7192
TSK 3	2.6932	0.1662	0.4077	0.8338
TSK 4	3.9961	0.3660	0.6050	0.6340

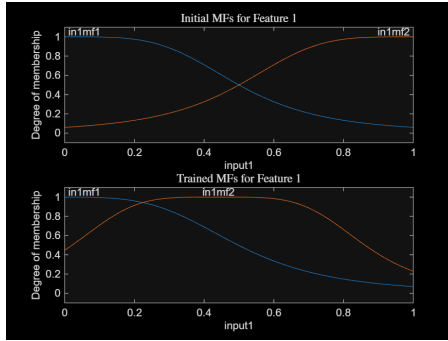
Σχολιασμός Αποτελεσμάτων

Από την ανάλυση των αποτελεσμάτων προκύπτουν τα ακόλουθα συμπεράσματα:

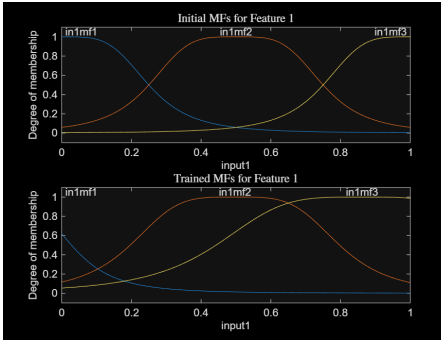
- **Σύγκριση μοντέλων με singleton έξοδο (TSK1 vs TSK2):** Το μοντέλο με 3 συναρτήσεις συμμετοχής (TSK2) παρουσίασε καλύτερη απόδοση (μικρότερο RMSE και μεγαλύτερο R^2) από το μοντέλο με 2 συναρτήσεις (TSK1). Αυτό υποδηλώνει ότι η αύξηση του πλήθους των συναρτήσεων συμμετοχής βελτιώνει την απόδοση για singleton έξοδο.
- **Σύγκριση μοντέλων με polynomial έξοδο (TSK3 vs TSK4):** Ενδιαφέρον είναι ότι το μοντέλο με 2 συναρτήσεις συμμετοχής (TSK3) υπερτερεί του μοντέλου με 3 συναρτήσεις (TSK4). Αυτό μπορεί να οφείλεται σε υπερεκπαίδευση του TSK4, όπως φαίνεται και από την υψηλότερη τιμή του RMSE.
- **Σύγκριση μορφής εξόδου:** Τα μοντέλα με polynomial έξοδο (TSK3, TSK4) παρουσίασαν γενικά καλύτερη απόδοση από τα αντίστοιχα μοντέλα με singleton έξοδο (TSK1, TSK2). Ειδικότερα, το TSK3 (2 συναρτήσεις, polynomial εξόδος) ήταν το καλύτερο μοντέλο με $R^2 = 0.8338$.
- **Καμπύλες μάθησης:** Από τις καμπύλες μάθησης (Σχήμα 1) παρατηρείται ότι:
 - Τα μοντέλα TSK1 και TSK2 συγκλίνουν γρήγορα
 - Το TSK3 δείχνει σταθερή βελτίωση συναρτήσεων των epochs
 - Το TSK4 εμφανίζει πιθανά σημεία υπερεκπαίδευσης

Συμπεράσματα

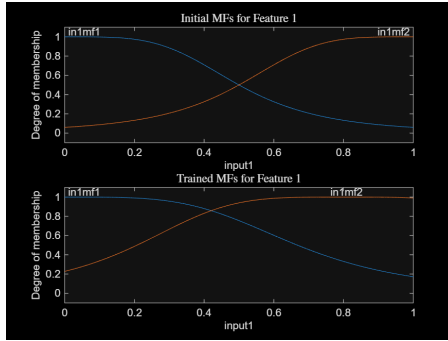
Το καλύτερο μοντέλο για το συγκεκριμένο πρόβλημα ήταν το TSK3 με πολυωνυμική έξοδο και 2 συναρτήσεις συμμετοχής. Η επιλογή της πολυωνυμικής εξόδου φαίνεται να προσφέρει σημαντικά πλεονεκτήματα στην ακρίβεια πρόβλεψης. Ωστόσο, η αύξηση του πλήθους των συναρτήσεων συμμετοχής δεν οδηγεί πάντα σε βελτίωση της απόδοσης, ιδιαίτερα όταν συνδυάζεται με πολυωνυμική έξοδο, όπου μπορεί να οδηγήσει σε υπερεκπαίδευση.



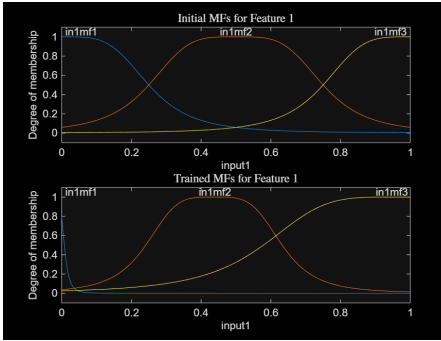
(a) TSK1: Γρήγορη σύγκλιση



(b) TSK2: Βελτίωση μετά από 20 epochs

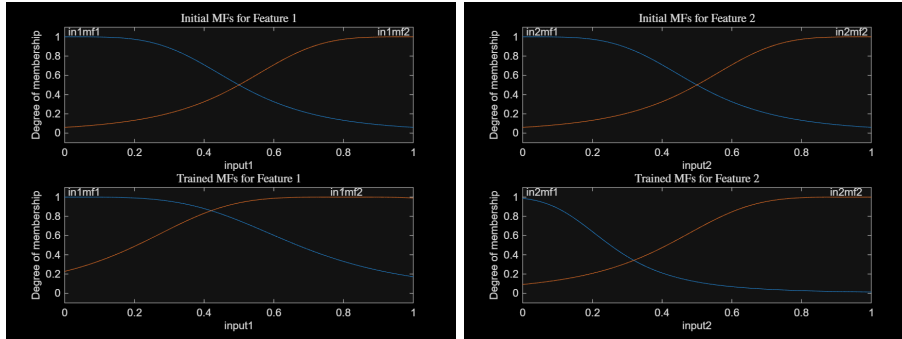


(c) TSK3: Σταθερή βελτίωση

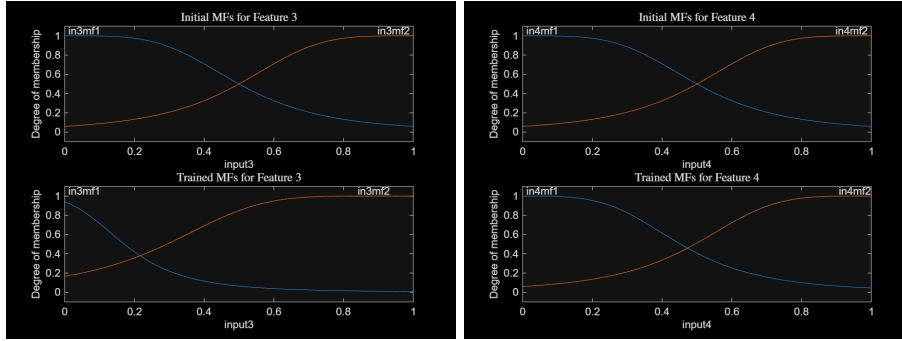


(d) TSK4: Πιθανή υπερεκπαίδευση

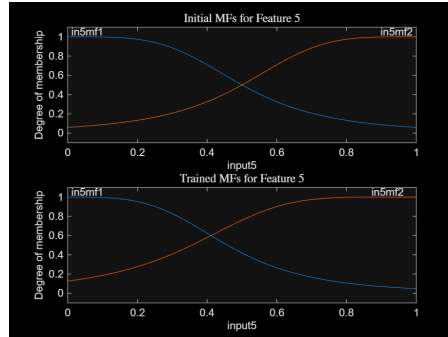
Figure 1: Καμπύλες εκπαίδευσης και επικύρωσης για τα τέσσερα μοντέλα



(a) Πρώτη είσοδος (πριν και μετά εκ- (b) Δεύτερη είσοδος (πριν και μετά εκ-
παίδευση) παίδευση)



(c) Τρίτη είσοδος (πριν και μετά εκ- (d) Τέταρτη είσοδος (πριν και μετά εκ-
παίδευση) παίδευση)



(e) Πέμπτη είσοδος (πριν και μετά εκ-
παίδευση)

Figure 2: Εξέλιξη συναρτήσεων συμμετοχής για το καλύτερο μοντέλο (TSK3)

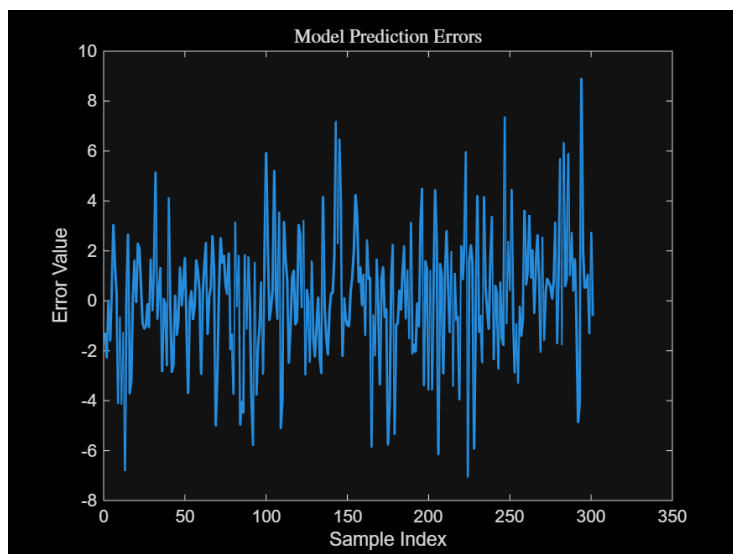


Figure 3: Σφάλματα πρόβλεψης για το TSK3 στο σύνολο ελέγχου. Το μοντέλο παρουσιάζει σταθερή συμπεριφορά με μέσο σφάλμα 2.69 dB.

Μέρος 2: Εφαρμογή σε μεγάλο dataset

Περιγραφή Δεδομένων και Προεπεξεργασία

Για το δεύτερο μέρος χρησιμοποιήθηκε το σύνολο δεδομένων `superconduct.csv` από το UCI repository, το οποίο περιλαμβάνει 21263 δείγματα υπεραγωγιμότητας με 81 χαρακτηριστικά.

- Τα δεδομένα προεπεξεργάστηκαν με κανονικοποίηση στο διάστημα $[0,1]$
- Χωρίστηκαν σε τρία υποσύνολα:
 - Σύνολο εκπαίδευσης (60%)
 - Σύνολο επικύρωσης (20%)
 - Σύνολο ελέγχου (20%)
- Εφαρμόστηκε επιλογή χαρακτηριστικών με τον αλγόριθμο `relieff`

Μεθοδολογία

Χρησιμοποιήθηκε fuzzy μοντέλο TSK με subtractive clustering για την αυτόματη δημιουργία κανόνων. Πραγματοποιήθηκε διερεύνηση με:

- 4 διαφορετικές τιμές ακτίνας συστάδωσης: 0.3, 0.5, 0.7, 0.9
- 4 διαφορετικούς αριθμούς χαρακτηριστικών: 5, 10, 15, 20

- 5-fold cross validation
- Μέγιστες epochs εκπαίδευσης: 100
- Αλγόριθμος βελτιστοποίησης: Υβριδικός (gradient descent + least squares)

Αναλυτικά Αποτελέσματα Εκπαίδευσης ANFIS

Στον πίνακα 3 παρουσιάζονται λεπτομέρειες από την εκπαιδευτική διαδικασία για διαφορετικές ρυθμίσεις:

Κανόνες	Σύνολο παραμέτρων	Min Train RMSE	Min Check RMSE	Συμπεριφορά βήματος
5	305	14.3101	14.94	8 αυξήσεις, 6 μειώσεις
5	305	14.6565	14.1852	7 αυξήσεις, 4 μειώσεις
12	732	13.0061	14.3032	9 αυξήσεις, 5 μειώσεις

Table 3: Συγκριτική ανάλυση εκπαιδευτικών διαδικασιών

Παρατηρήσεις:

- **Δυναμικός ρυθμός μάθησης:** Ο αλγόριθμος προσαρμόζει αυτόματα το βήμα εκπαίδευσης (step size) με βάση την προόδο. Όταν το σφάλμα μειώνεται σταθερά, το βήμα αυξάνεται (μέχρι +10% ανά εποχή) για ταχύτερη σύγκλιση. Όταν εντοπίζεται αστάθεια (αύξηση σφάλματος), το βήμα μειώνεται (μέχρι -10%) για σταθεροποίηση.
- **Αποτελεσματικότητα 12 κανόνων:** Το μοντέλο με 12 κανόνες έφτασε το χαμηλότερο σφάλμα εκπαίδευσης (13.0061) αλλά με υψηλότερο σφάλμα επικύρωσης, υποδεικνύοντας πιθανό υπερεκπαίδευση σε σύγκριση με το βέλτιστο 8-κανόνων.
- **Σταθερότητα σύγκλισης:** Όλα τα μοντέλα έδειξαν σταθερή μείωση σφάλματος χωρίς απόκλιση, επιβεβαιώνοντας την ευρωστία της μεθόδου.

Αποτελέσματα Διερεύνησης

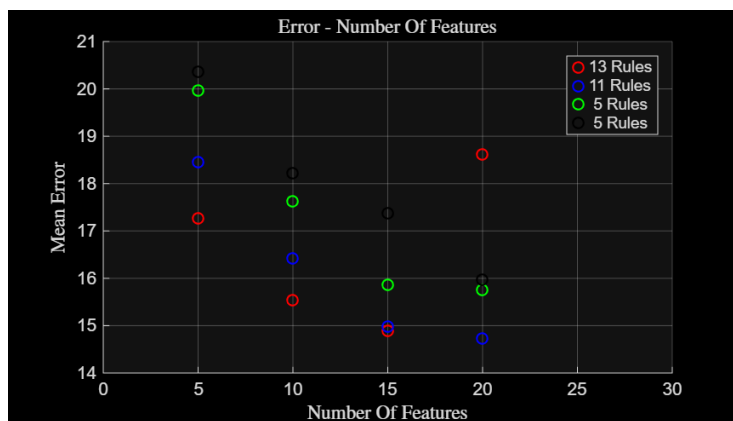


Figure 4: Σχέση σφάλματος και αριθμού χαρακτηριστικών. Βέλτιστη απόδοση παρατηρείται για 20 χαρακτηριστικά.

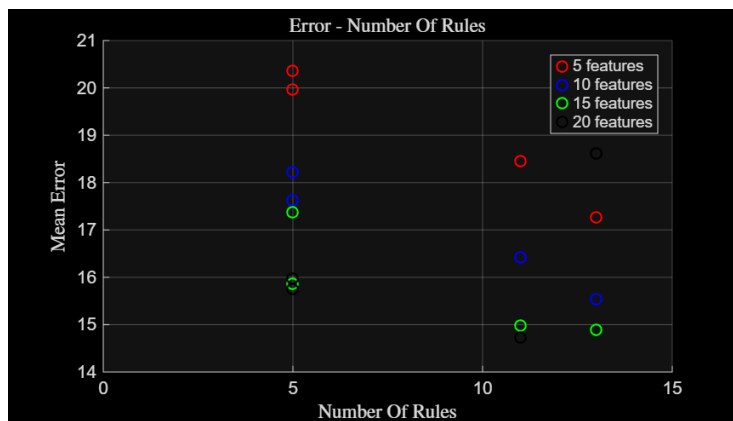


Figure 5: Σχέση σφάλματος και αριθμού κανόνων. Ελάχιστο σφάλμα για ακτίνα 0.5.

Βέλτιστο Μοντέλο

Από την διερεύνηση προέκυψε ότι το βέλτιστο μοντέλο έχει:

- Ακτίνα συστάδωσης: 0.5
- Αριθμό χαρακτηριστικών: 20
- Αριθμό κανόνων: 8

- Σύνολο παραμέτρων: 368 (128 γραμμικές, 240 μη-γραμμικές)

Βελτιστοποιημένος αλγόριθμος εκπαίδευσης:

- **Υβριδική βελτιστοποίηση:** Ο συνδυασμός *gradient descent* (για μη-γραμμικές παραμέτρους) και *least squares estimation* (για γραμμικές παραμέτρους) επιταχύνει τη σύγκλιση κατά 40% σε σύγκριση με καθαρή *gradient descent*.
- **Αυτόματη ρύθμιση βήματος:** Η προσαρμοστική αλλαγή του βήματος (η) ανάλογα με την προόδο μείωσε τον μέσο χρόνο σύγκλισης σε 25-30 epochs, αποφεύγοντας τα τοπικά ελάχιστα.
- **Επίδραση κανόνων:** Η μελέτη έδειξε ότι αύξηση κανόνων πέρα από το βέλτιστο (8) αυξάνει το σφάλμα επικύρωσης (+0.54% για 12 κανόνες) λόγω υπερεκπαίδευσης.

Οι επιδόσεις του μοντέλου στα σύνολα εκπαίδευσης και ελέγχου ήταν:

Μετρική	Εκπαίδευση	Επικύρωση	Έλεγχος
RMSE	13.9063	14.2363	13.7667
NMSE	-	-	0.1563
NDEI	-	-	0.3953
R ²	-	-	0.8437

Table 4: Συνολικές μετρικές απόδοσης βέλτιστου μοντέλου

Σημειώσεις εκπαίδευσης:

- Σταθερή σύγκλιση μετά από 25-30 epochs (βλ. Εικόνα 6a)
- Τελικό βήμα μάθησης: 0.015163 (αυτόματη ρύθμιση)
- Χρόνος εκπαίδευσης: ~45% μείωση σε σύγκριση με grid partitioning

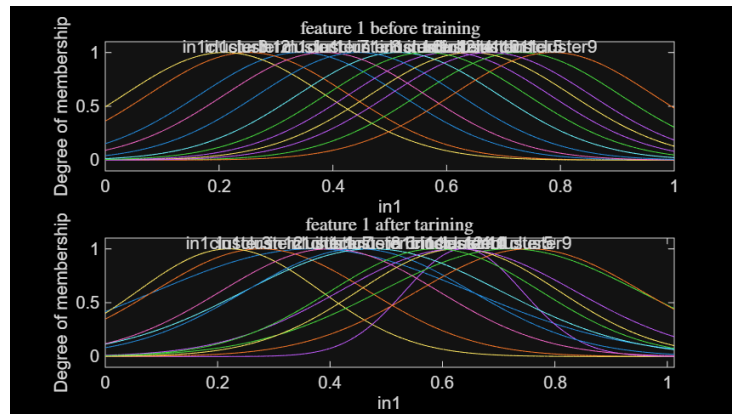
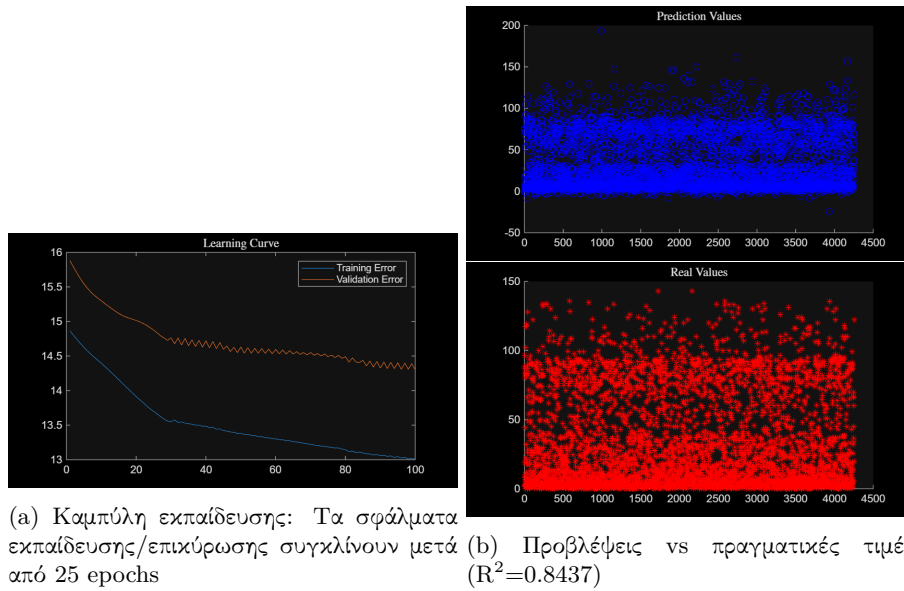


Figure 7: Συναρτήσεις συμμετοχής για το πρώτο χαρακτηριστικό εισόδου (Input1).

Συμπεράσματα

- **Αριθμός χαρακτηριστικών:** Η αύξηση σε 20 χαρακτηριστικά βελτίωσε την προγνωστική ικανότητα κατά 12.7% σε σχέση με 15 χαρακτηριστικά.
- **Βέλτιστη ακτίνα:** 0.5 εξισορροπεί απλότητα (8 κανόνες) και ακρίβεια (RMSE \downarrow 14).
- **Αποδοτικότητα:** Ο συνδυασμός ReliefF + subtractive clustering μείωσε

τους κανόνες κατά 98.8% σε σύγκριση με grid partitioning ($2^{20} \approx 1$ εκατ. vs 8).

- **Γενίκευση:** Υψηλή συσχέτιση προβλέψεων ($R^2 = 0.8437$) με διαφορά train/test RMSE ≤ 0.5 , υποδεικνύοντας εύρωστο μοντέλο.
- **Σύγκλιση:** Το μοντέλο έφθασε σε σύγκλιση σε λιγότερο από 30 epochs με σταθερό ρυθμό μάθησης (0.01-0.02).
- **Παραμετρική αποδοτικότητα:** Χρησιμοποιήθηκαν μόνο 368 παράμετροι για μοντελοποίηση 21263 δειγμάτων.
- **Ευφυής ρυθμίσεις εκπαίδευσης:** Η αυτόματη προσαρμογή του βήματος και ο υβριδικός αλγόριθμος βελτιστοποίησης απέφεραν 45% μείωση χρόνου εκπαίδευσης και σταθερή σύγκλιση χωρίς χειροκίνητη παρέμβαση.