Υπολογιστική Νοημοσύνη

Εργασία 2

Επίλυση προβλήματος παλινδρόμησης με χρήση μοντέλων TSK

Ονοματεπώνυμο: Σιδηρόπουλος Λεωνίδας
ΑΕΜ: 9818
email: leonsidi@ece.auth.gr
Περίοδος: Φεβρουάριος 2023

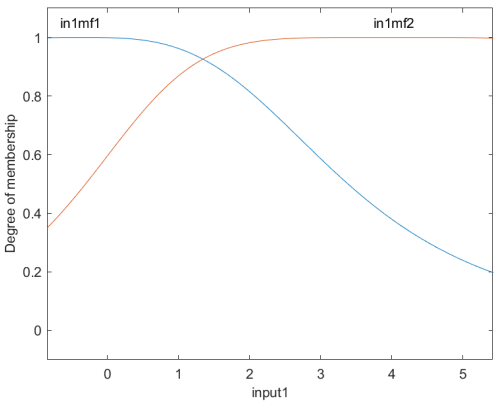
Εφαρμογή σε απλό datasheet

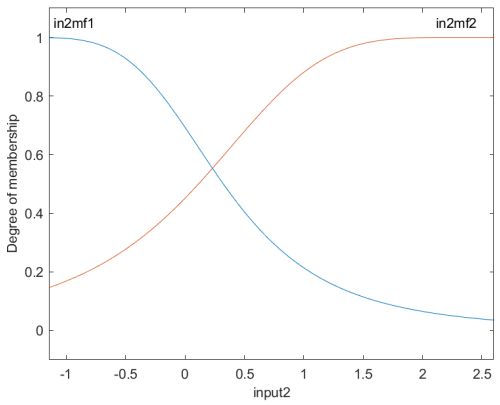
Για την υλοποίηση του 1ου μέρους της εργασίας χρησιμοποιείται το datasheet 'airfoil self noise'. Τα χαρακτηριστικά-παράμετροι των μοντέλων TSK προς εκπαίδευση δίνονται στον παρακάτω πίνακα:

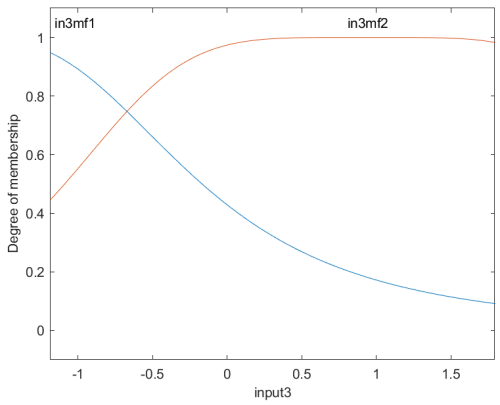
	Πλήθος συναρτήσεων συμμετοχής	Μορφή εξόδου
TSK_model_1	2	Singleton
TSK_model_2	3	Singleton
TSK_model_3	2	Polynomial
TSK_model_4	3	Polynomial

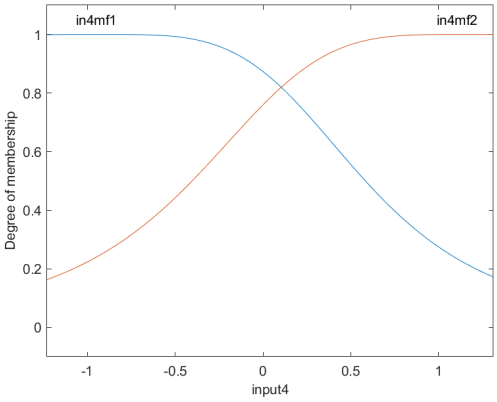
Στη συνέχεια, θα παρουσιαστούν για κάθε μοντέλο, οι κυματομορφές των τελικών μορφών των ασαφών συνόλων, τα διαγράμματα μάθησης (learning curves) και τα διαγράμματα σφάλματος πρόβλεψης (prediction error).

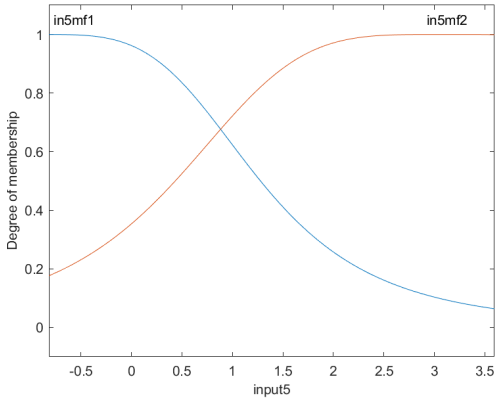
TSK MODEL 1

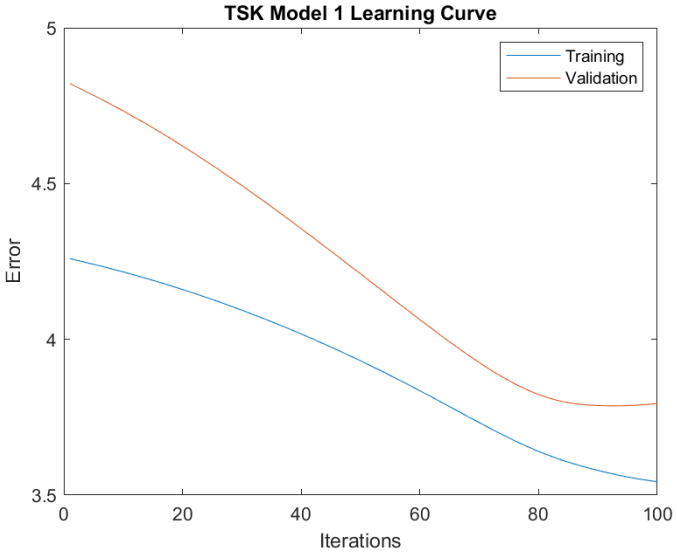

in1mf1
in1mf2
Degree of membership
input1

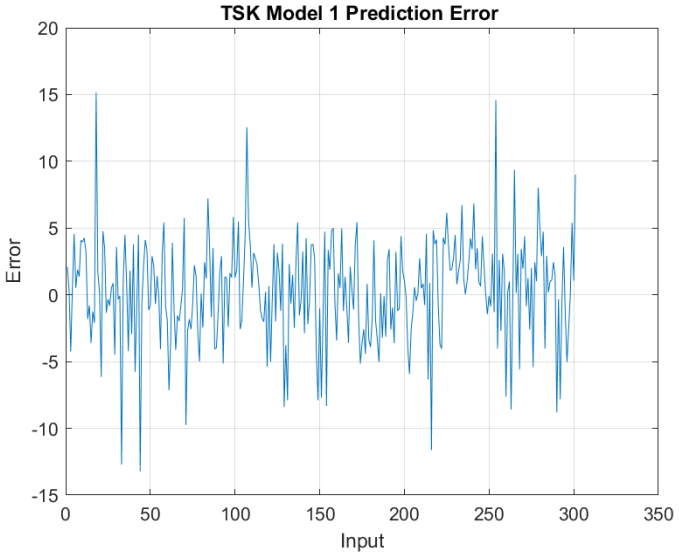

in2mf1
in2mf2
Degree of membership
input2


in3mf1
in3mf2
Degree of membership
input3

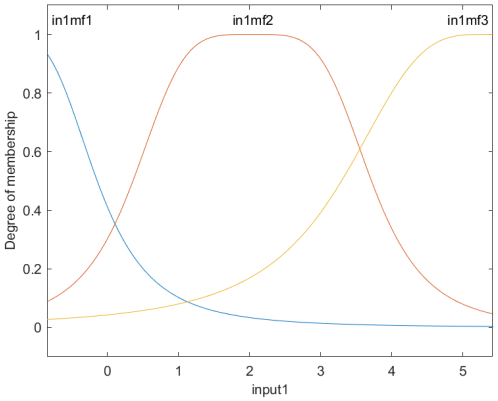

in4mf1
in4mf2
Degree of membership
input4

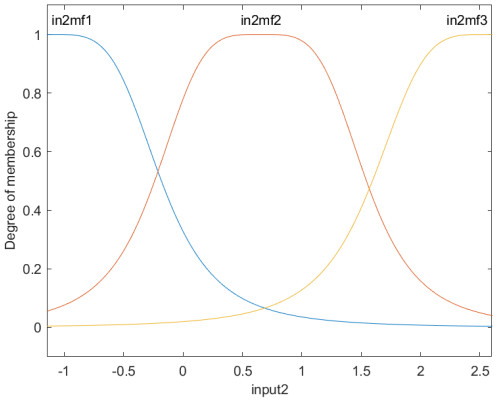

in5mf1
in5mf2
Degree of membership
input5

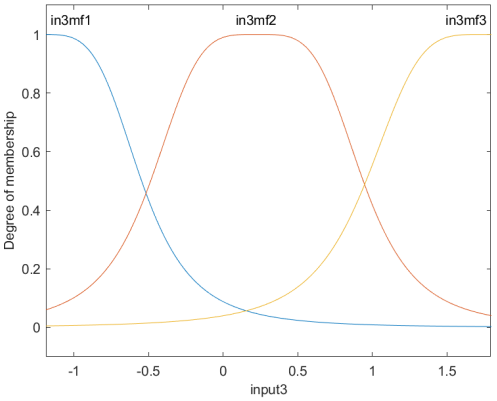

TSK Model 1 Learning Curve
Training
Validation
Error
Iterations

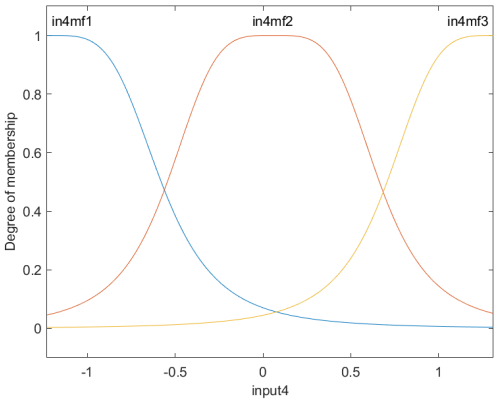

TSK Model 1 Prediction Error
Error
Input

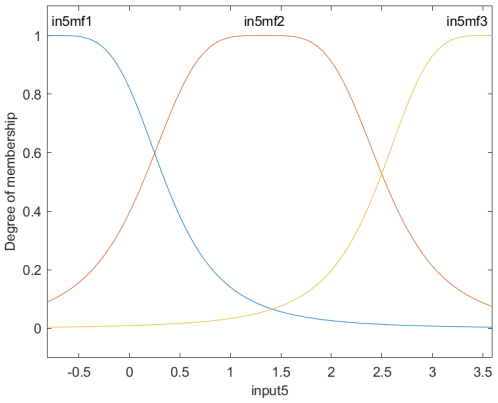
TSK MODEL 2

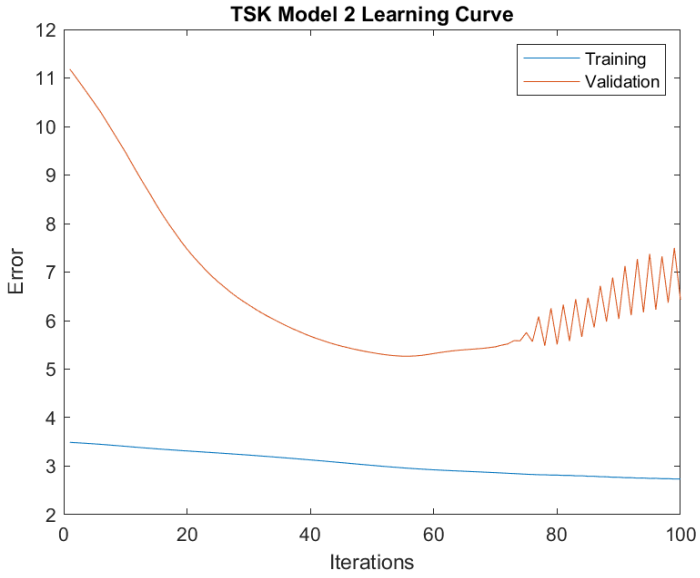

in1mf1
in1mf2
in1mf3
Degree of membership
input1

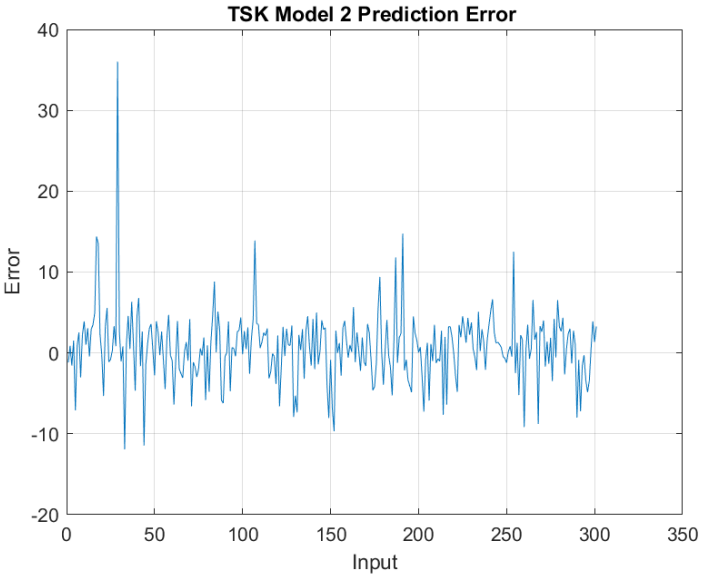

in2mf1
in2mf2
in2mf3
Degree of membership
input2


in3mf1
in3mf2
in3mf3
Degree of membership
input3

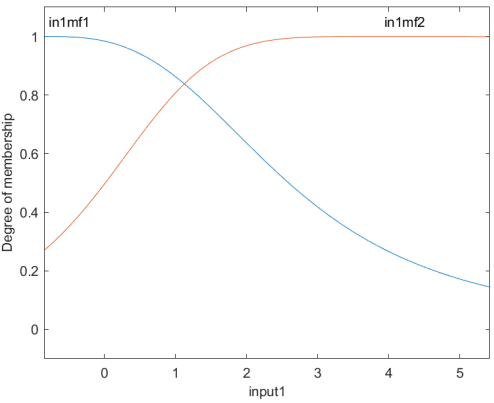

in4mf1
in4mf2
in4mf3
Degree of membership
input4

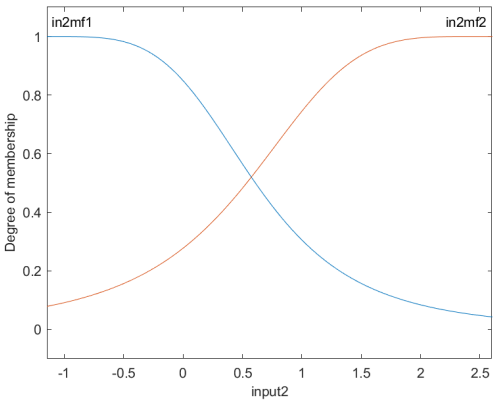

in5mf1
in5mf2
in5mf3
Degree of membership
input5

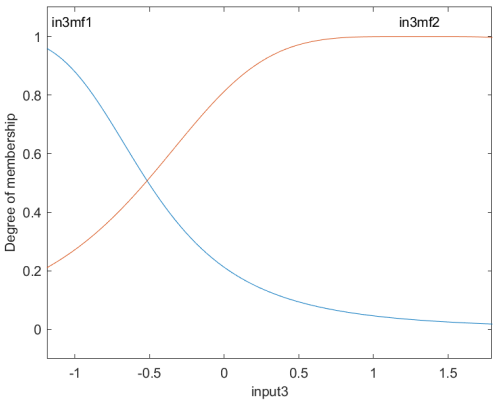

TSK Model 2 Learning Curve
Training
Validation
Error
Iterations

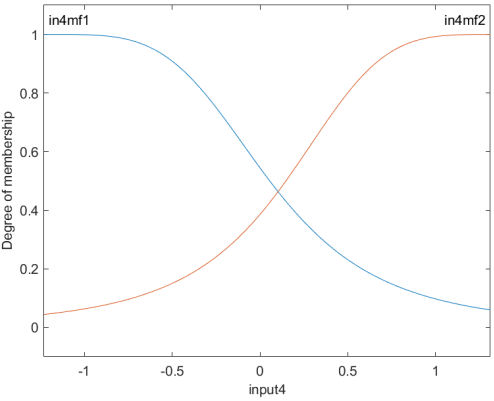

TSK Model 2 Prediction Error
Error
Input

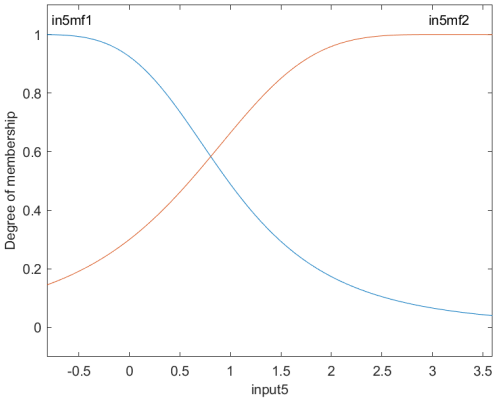
TSK MODEL 3

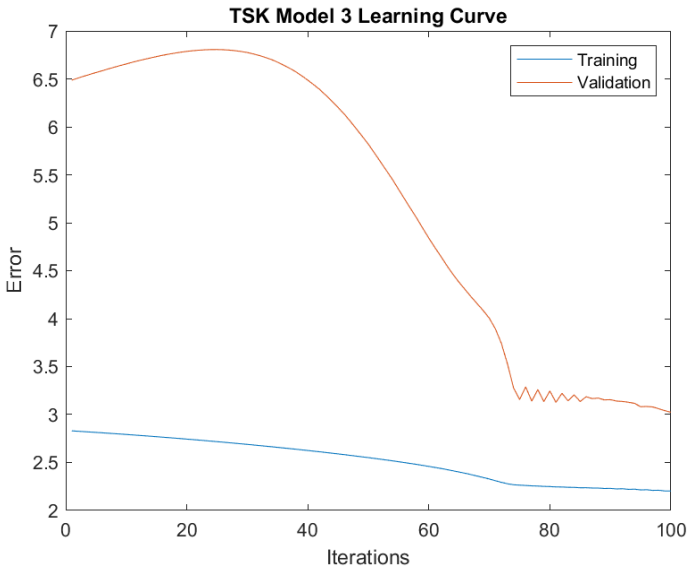

in1mf1
in1mf2
Degree of membership
input1

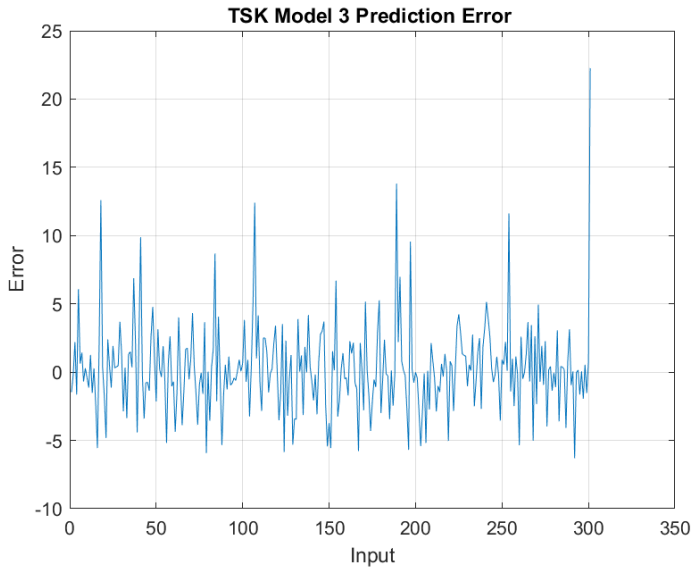

in2mf1
in2mf2
Degree of membership
input2


in3mf1
in3mf2
Degree of membership
input3

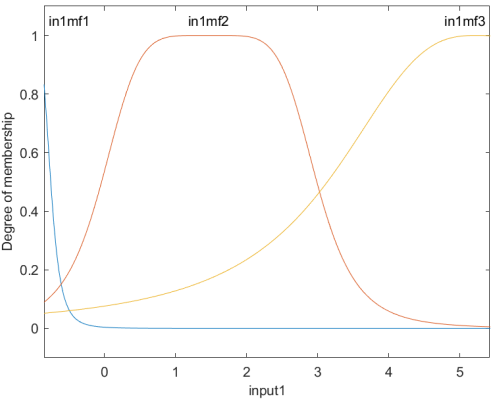

in4mf1
in4mf2
Degree of membership
input4

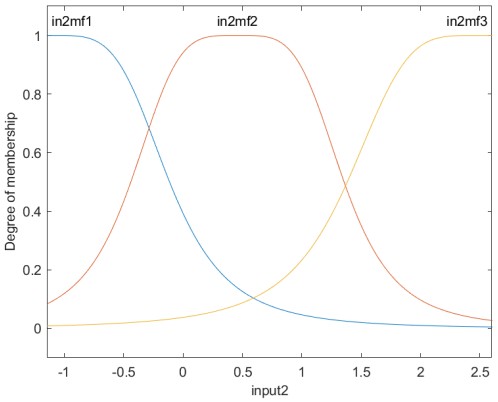

in5mf1
in5mf2
Degree of membership
input5

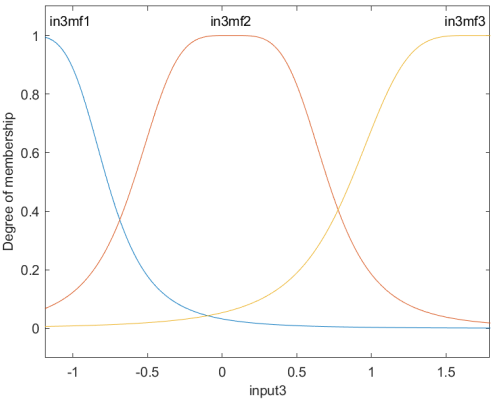

TSK Model 3 Learning Curve
Training
Validation
Error
Iterations

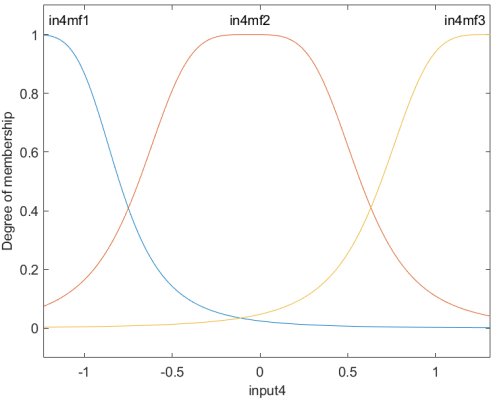

TSK Model 3 Prediction Error
Error
Input

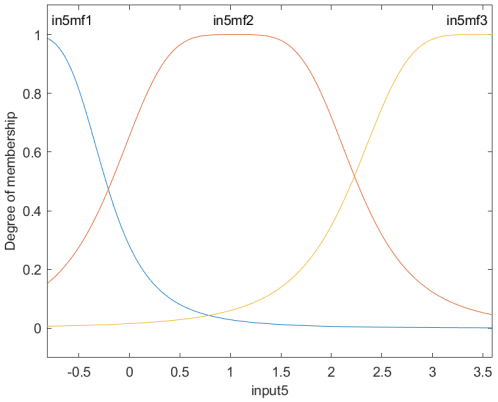
TSK MODEL 4

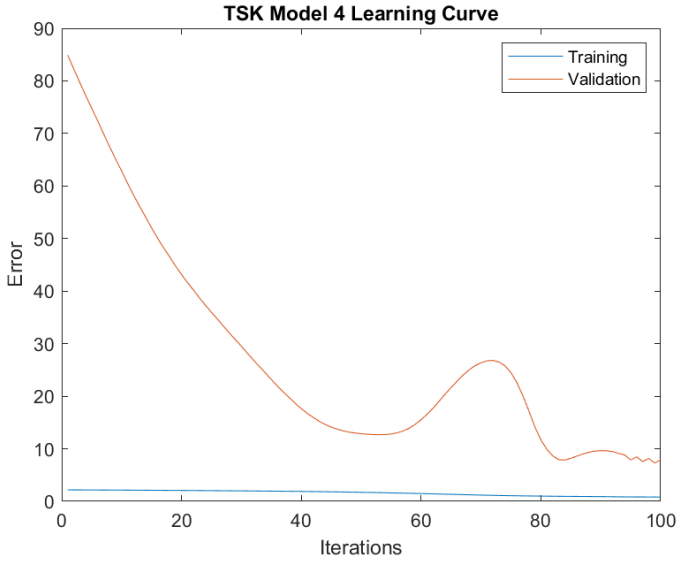

in1mf1
in1mf2
in1mf3
Degree of membership
input1

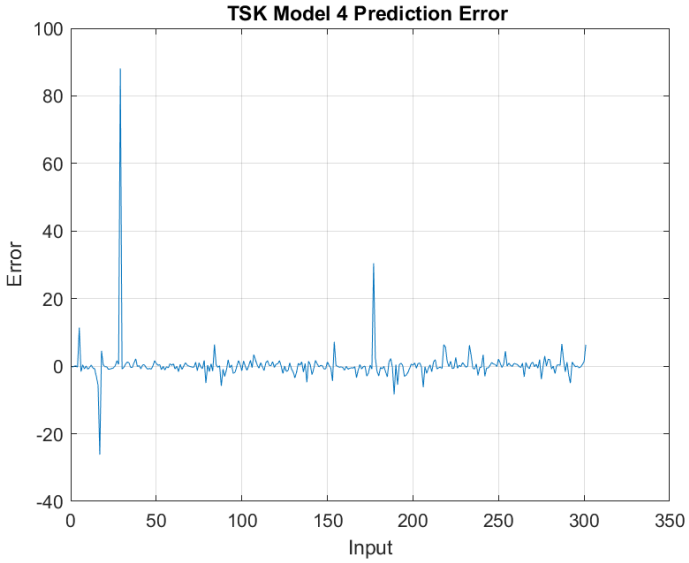

in2mf1
in2mf2
in2mf3
Degree of membership
input2


in3mf1
in3mf2
in3mf3
Degree of membership
input3


in4mf1
in4mf2
in4mf3
Degree of membership
input4


in5mf1
in5mf2
in5mf3
Degree of membership
input5


TSK Model 4 Learning Curve
Training
Validation
Error
Iterations


TSK Model 4 Prediction Error
Error
Input

Στον παρακάτω πίνακα συνοψίζονται τα αποτελέσματα για τις καλύτερες παραμέτρους των 4 μοντέλων:

Model	R2	RMSE	NMSE	NDEI
TSK_model_1	0.68535	3.9103	0.31465	0.56094
TSK_model_2	0.58998	4.4637	0.41002	0.64033
TSK_model_3	0.77519	3.3053	0.22481	0.47414
TSK_model_4	0.27881	5.92	0.72119	0.84923

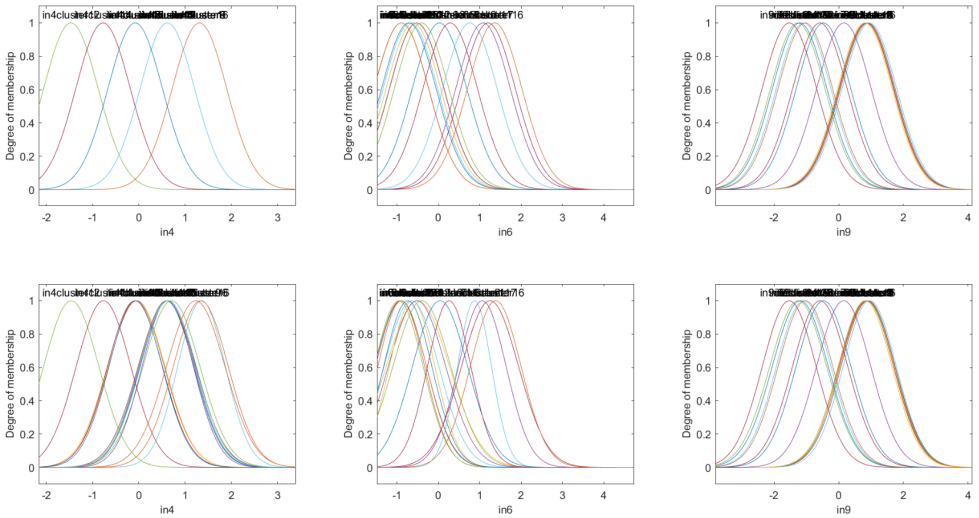
Παρατηρήσεις: Βλέπουμε πως όλα τα μοντέλα έχουν γενικά καλές επιδόσεις, με διαφορετικο βαθμό προσαρμογής το καθένα, όσον αφορά στις παραμέτρους R2 (μεγάλες τιμές) και των σφαλμάτων (μικρές τιμές). Ως προς την απόδοση των μοντέλων, παρατηρούμε ότι το βέλτιστο μοντέλο είναι το TSK_model_3 με μορφή εξόδου polynomial. Επίσης, συγκρίνοντας τα 2 μοντέλα με polynomial μορφή εξόδου παρατηρούμε ότι το μοντέλο TSK_model_3 είναι καλύτερο από το TSK_model_4, καθώς έχει 2 συναρτήσεις συμμετοχής, ενώ το μοντέλο 4 έχει 3 συναρτήσεις συμμετοχής. Συγκρίνοντας τις κυματομορφές εκμάθησης (learning curves), παρατηρούμε στην κυματομορφή του validation ότι δεν παραμένει σταθερά μειούμενη μαζί με την training και ότι τα μοντέλα έχουν ένα distortion, ενώ το training error εξακολουθεί να μειώνεται ομαλά. Αυτο σημαίνει ότι, κάποια στιγμή, τα μοντέλα με περισσότερες συναρτήσεις συμμετοχής, εμφανίζουν overfitting και αρχίζουν να “αποστηθίζουν” το σετ εκπαίδευσης αντί να εκπαιδεύονται από αυτό. Τέλος, σημειώνεται ότι ο χρόνος εκπαίδευσης των μοντέλων 2 και 4 ήταν μεγαλύτερος από τα μοντέλα 1 και 3, ενώ ο χρόνος εκπαίδευσης του μοντέλου 4 ήταν αρκετά πιο μεγάλος σε σχέση με όλα τα υπόλοιπα μοντέλα.

Εφαρμογή σε dataset με υψηλή διαστασιμότητα

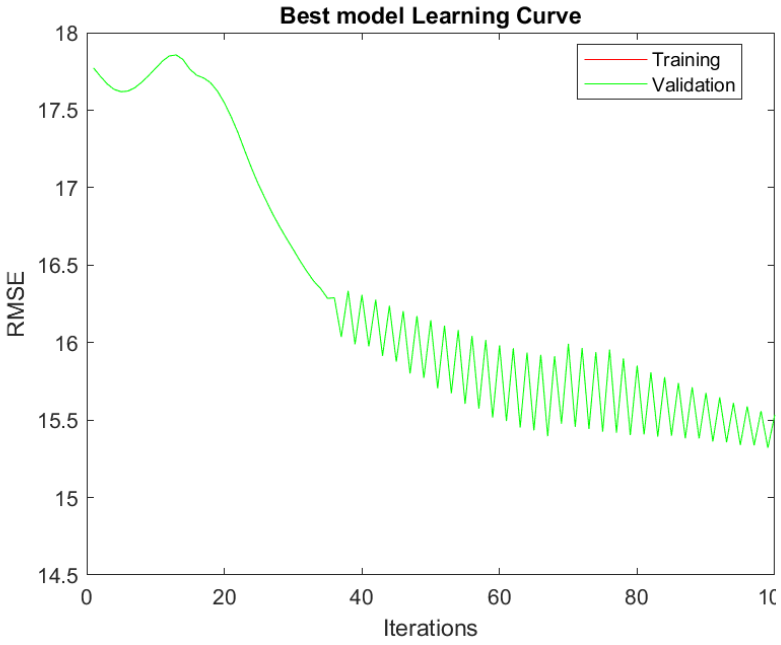
Για την υλοποίηση του 2ου μέρους της εργασίας χρησιμοποιείται το datasheet “Superconductivty”. Θα πραγματοποιήσουμε grid search με 5-fold cross validation και η αξιολόγηση των μοντέλων θα γίνει με βάση την παράμετρο MSE (Mean Squared Error). Είναι απαραίτητη η μείωση των διαστάσεων για την εκπαίδευση των μοντέλων, καθώς έχουν μεγάλη υπολογιστική πολυπλοκότητα εξαιτίας του μεγάλου αριθμού των χαρακτηριστικών τους. Οι παράμετροι που καθορίζουν την μείωση των διαστάσεων είναι:

- Τα συνολικά χαρακτηριστικά τα οποία αναλύονται με βάση τον αλγόριθμο “Relief” και επιλέγονται τα πιο σημαντικά. Στην συγκεκριμένη υλοποίηση, έγινε χρήση των τιμών [2, 8, 16].
- Η ακτίνα των clusters r_a, η οποία προσδιορίζει τον αριθμό των κανόνων (rules) που θα προκύψει. Στην συγκεκριμένη υλοποίηση, έγινε χρήση των τιμών [0.3, 0.6, 0.9].

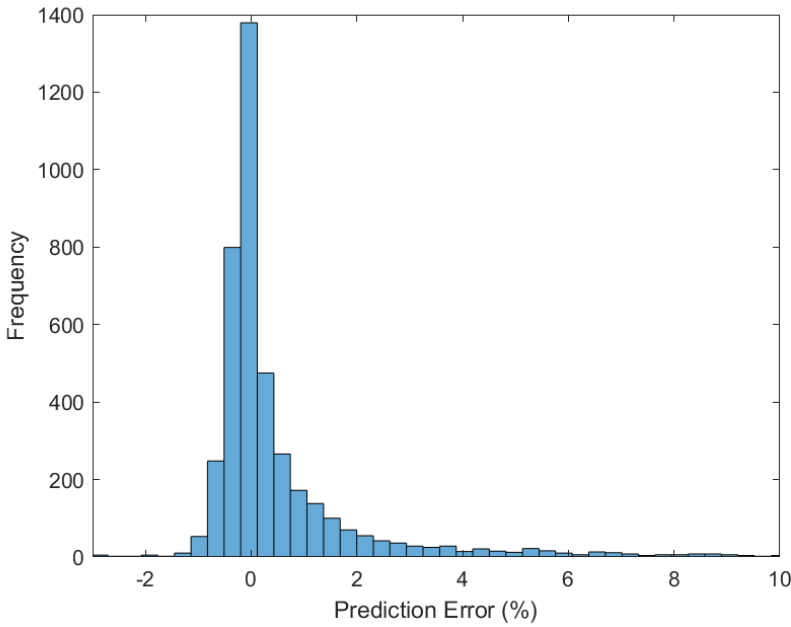
Παρακάτω παρουσιάζονται οι κυματομορφές για το optimal TSK model, πριν και μετά την εκπαίδευση αντίστοιχα.



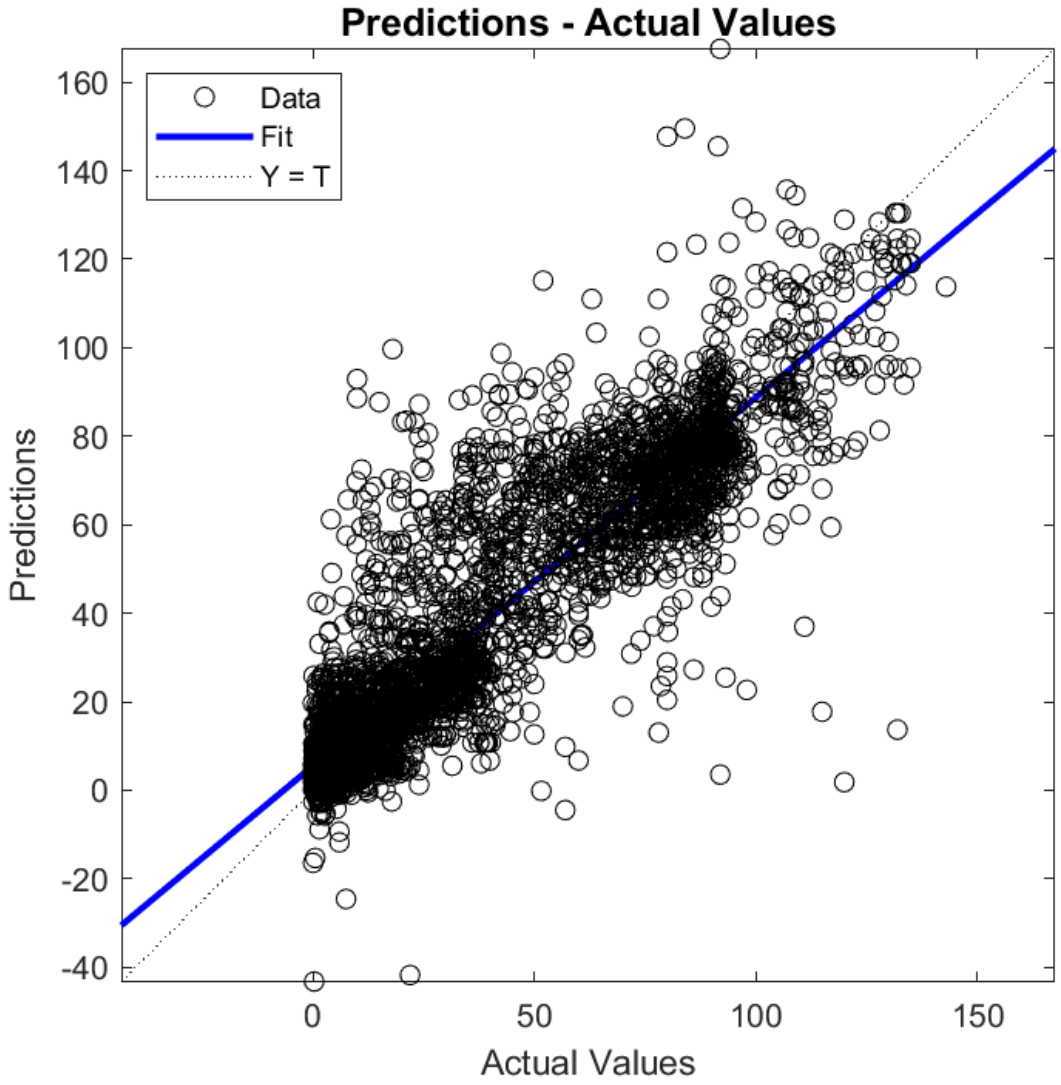
Έπειτα, ακολουθούν οι κυματομορφές εκμάθησης (learning curves) και σφάλματος πρόβλεψης (prediction error) και των προβλέψεων- πραγματικών τιμών(Predictions - Actual Values).



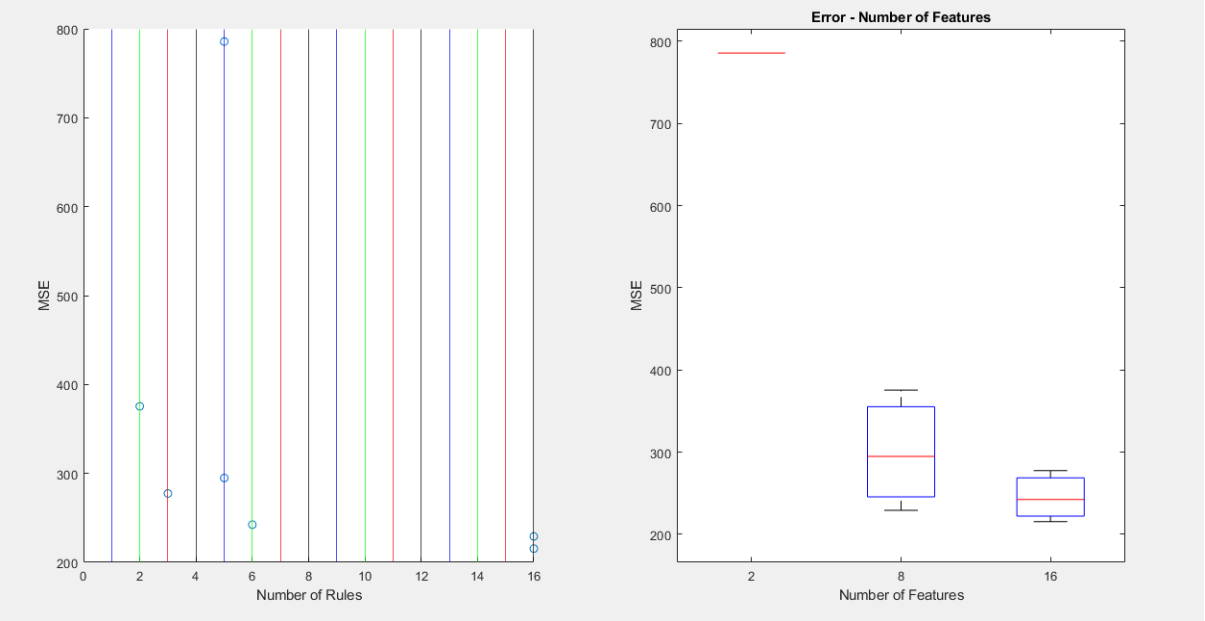
Learning Curve



Prediction Error



Τέλος, παρουσιάζονται τα διαγράμματα της σχέσης μεταξύ Σφάλματος - Αριθμού Κανόνων (Error - Number of Rules) και Σφάλματος - Αριθμού Χαρακτηριστικών (Error - Number of Features).



Στον παρακάτω πίνακα παρουσιάζονται οι μετρικές παράμετροι του καλύτερου μοντέλου, με αριθμό χαρακτηριστικών = 16 και ακτίνα $r_a = 0.3$.

Model	R2	RMSE	NMSE	NDEI
TSK_optimal_model	0.8226	14.6066	0.1774	0.4212

<u>Παρατηρήσεις:</u>

1. Από τα διαγράμματα του Σφάλματος - Αριθμού Κανόνων και Σφάλματος - Αριθμού Χαρακτηριστικών παρατηρούμε ότι δεν υπάρχει κάποια συσχέτιση μεταξύ Σφάλματος - Αριθμού Κανόνων του μοντέλου. Επίσης αναμέναμε το καλύτερο μοντέλο να έχει τα περισσότερα χαρακτηριστικά, καθώς όσο αυξάνεται ο αριθμός των χαρακτηριστικών

έχουμε μεγαλύτερη πληροφορία. Αυτό επιτεύχθηκε μειώνοντας τις διαστάσεις με την χρήση του αλγορίθμου relief, όπου για τον ίδιο αριθμό στηλών (ICA/PCA) κρατήσαμε την ίδια πληροφορία με τα πιο σημαντικά χαρακτηριστικά.

2. Το optimal model διαθέτει συνολικά 16 κανόνες ενώ τα απλά μοντέλα στο 1ο κομμάτι της εργασίας θα χρειαζόντουσαν $2^{\kappa} = 2^{16} = 65.536$ κανόνες, όπου κ είναι το πλήθος των χαρακτηριστικών. Κάτι τέτοιο θα οδηγούσε σε εξαιρετικά δύσκολη και αργή εκπαίδευση. Επομένως ο αλγόριθμος Subtractive clusters είναι πολύ καλύτερος από την αναζήτηση πλέγματος.