



Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης
Πολυτεχνική Σχολή
Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
Τομέας Ηλεκτρονικής και Υπολογιστών

Επίλυση Προβλήματος παλινδρόμησης με χρήση μοντέλων TSK

Εργασία για το μάθημα
Υπολογιστική Νοημοσύνη
του
Μηνά Κοσμίδα
ΑΕΜ: 9008

Διδάσκοντες: Ιωάννης Θεοχάρης
Καθηγητής

Χρήστος Χαδουλός
Μεταπτυχιακός Φοιτητής

Φεβρουάριος 2023

Περιεχόμενα

1 Εφαρμογή σε απλό Dataset	3
1.1 Προδιαγραφές	3
1.2 Μειτρικές Αξιολόγησης	4
1.3 Εκπαίδευση	5
1.3.1 Μοντέλο 1 (2 MFs, Singleton)	5
1.3.2 Μοντέλο 2 (3 MFs, Singleton)	7
1.3.3 Μοντέλο 3 (2 MFs, Polynomial)	9
1.3.4 Μοντέλο 4 (3 MFs, Polynomial)	11
1.4 Αξιολόγηση	13
2 Εφαρμογή σε δεδομένα υψηλής διαστασιμότητας	14
2.1 Προδιαγραφές	14
2.2 Grid Search	15
2.3 Βέλτιστο μοντέλο	16

Εισαγωγή

Η παρούσα εργασία εκπονήθηκε στα πλαίσια του μαθήματος Υπολογιστική Νοημοσύνη κατά το χειμερινό εξάμηνο του έτους 2022-2023. Η ανάπτυξη αυτής πραγματοποιήθηκε στο περιβάλλον προγραμματισμού MATLAB (έκδοση R2022b) και αποτελεί την τρίτη από μια σειρά τεσσάρων ασκήσεων σύμφωνα με την κατηγορία ΠΠΣ. **Η συγγραφή της αναφοράς βασίστηκε στα δεδομένα της τελευταίας εκτέλεσης των προγραμμάτων πριν την παράδοσή της.**

Περιγραφή Προβλήματος

Στόχος της εργασίας είναι να διερευνηθεί η ικανότητα των μοντέλων TSK στη μοντελοποίηση πολυμεταβλητών, μη γραμμικών συναρτήσεων. Συγκεκριμένα, επιλέγονται δύο σύνολα δεδομένων από το UCI repository με σκοπό την εκτίμηση της μεταβλητής στόχου από τα διαθέσιμα δεδομένα, με χρήση ασαφών νευρωνικών μοντέλων. Το πρώτο σύνολο δεδομένων θα χρησιμοποιηθεί για μια απλή διερεύνηση της διαδικασίας εκπαίδευσης και αξιολόγησης μοντέλων αυτού του είδους, καθώς και για μια επίδειξη τρόπων ανάλυσης και ερμηνείας των αποτελεσμάτων. Το δεύτερο, πολυπλοκότερο, σύνολο δεδομένων θα χρησιμοποιηθεί για μια πληρέστερη διαδικασία μοντελοποίησης, η οποία θα περιλαμβάνει μεταξύ άλλων προ-επεξεργαστικά βήματα όπως επιλογή χαρακτηριστικών (feature selection), καθώς και μεθόδους βελτιστοποίησης των μοντέλων μέσω της διασταυρωμένης επικύρωσης (cross validation).

Κεφάλαιο 1

Εφαρμογή σε απλό Dataset

1.1 Προδιαγραφές

Για το πρώτο κομμάτι της εργασίας, καλούμαστε να εκπαιδεύσουμε τέσσερα μοντέλα με διαφορετικά χαρακτηριστικά όπως περιγράφονται από τον Πίνακα 1.1. Επιλέγεται από το UCI repository το Airfoil Self-Noise dataset, το οποίο περιλαμβάνει 1503 δείγματα και 6 χαρακτηριστικά, το τελευταίο εκ των οποίων είναι το εξαρτώμενο χαρακτηριστικό ή target feature.

Αρχικά είναι αναγκαίος ο διαχωρισμός των δεδομένων σε τρία μη επικαλυπτόμενα υποσύνολα, *Εκπαίδευσης* (D_{trn}), *Επικύρωσης* (D_{val}), και *Ελέγχου* (D_{test}). Όπως υποδεικνύουν και τα ονόματα των συνόλων, το πρώτο χρησιμοποιείται για την εκπαίδευση, το δεύτερο για αποφυγή του φαινομένου υπερεκπαίδευσης και το τρίτο για έλεγχο της απόδοσης του τελικού μοντέλου.

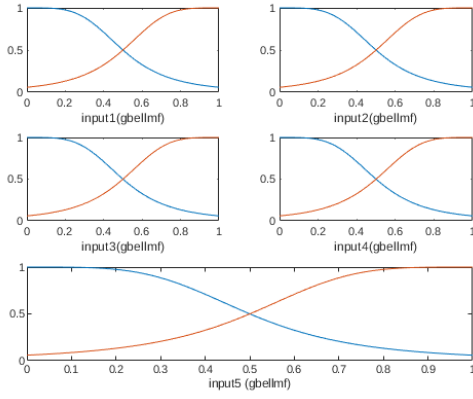
Από την θεωρία ο διαχωρισμός γίνεται σε 60%, 20%, και 20%, του αρχικού συνόλου δεδομένων, στο κάθε υποσύνολο αντίστοιχα.

Περιγραφή Μοντέλων		
Μοντέλο	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

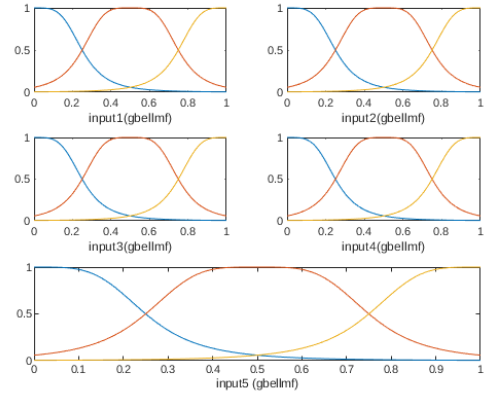
Πίνακας 1.1: Ταξινόμηση μοντέλων προς εκπαίδευση

Και τα τέσσερα μοντέλα εκπαιδεύονται με την υβριδική μέθοδο, σύμφωνα με την οποία οι παράμετροι των συναρτήσεων συμμετοχής βελτιστοποιούνται μέσω της μεθόδου της οπισθοδιάδοσης (backpropagation algorithm), ενώ οι παράμετροι της πολυωνμικής συνάρτησης εξόδου βελτιστοποιούνται μέσω της μεθόδου των ελαχίστων τετραγώνων (Least Squares).

Ο τύπος των συναρτήσεων συμμετοχής ορίζεται να είναι bell-shaped και η αρχικοποίησή τους γίνεται με τρόπο ώστε τα διαδοχικά ασαφή σύνολα να παρουσιάζουν, σε κάθε είσοδο, βαθμό επικάλυψης περίπου 0.5 σύμφωνα με την εικόνα 1.1



(α) Είσοδοι 2 Συναρτήσεων Συμμετοχής



(β) Είσοδοι 3 Συναρτήσεων Συμμετοχής

Σχήμα 1.1: Αρχικοποίηση Συναρτήσεων Συμμετοχής

1.2 Μετρικές Αξιολόγησης

Για την ακρίβεια της εκτίμησης της πραγματικής συνάρτησης από καθένα από τα παραπάνω μοντέλα, θα χρησιμοποιηθούν οι εξής δείκτες απόδοσης:

1. MSE: το μέσο τετραγωνικό σφάλμα μεταξύ της εξόδου του μοντέλου και της πραγματικής εξόδου

$$MSE(\theta) = \sigma_e^2 = \frac{1}{N} \sum_{i=1}^N e_i^2 = \sum_{i=1}^N (f(\mathbf{x}_i) - \hat{f}(\mathbf{x}_i))^2 \quad (1.1)$$

εναλλακτικά μπορεί να χρησιμοποιηθεί ο δείκτης RMSE, ο οποίος είναι η τετραγωνική ρίζα του MSE:

$$RMSE(\theta) = \sqrt{MSE(\theta)} \quad (1.2)$$

2. Συντελεστής προσδιορισμού R^2 : Ο συντελεστής αυτός, παρέχει πληροφορία για το ποσοστό της διακύμανσης της πραγματικής εξόδου που "εξηγείται" από το μοντέλο μας.

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (1.3)$$

όπου y είναι η πραγματική έξοδος του συστήματος, \hat{y} η εκτίμηση του μοντέλου και \bar{y} η μέση τιμή της πραγματικής εξόδου.

3. Τέλος, μπορούν να χρησιμοποιηθούν και οι δείκτες NMSE (*Normalized Mean Square Error*) και NDEI (*Non-Dimensional Error Index*)

$$NMSE = \frac{\sigma_e^2}{\sigma_x^2} = \frac{\sum_{k=1}^N (y_k - \hat{y}_k)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (1.4)$$

$$NDEI = \sqrt{NMSE} = \frac{\sigma_e}{\sigma_x} \quad (1.5)$$

1.3 Εκπαίδευση

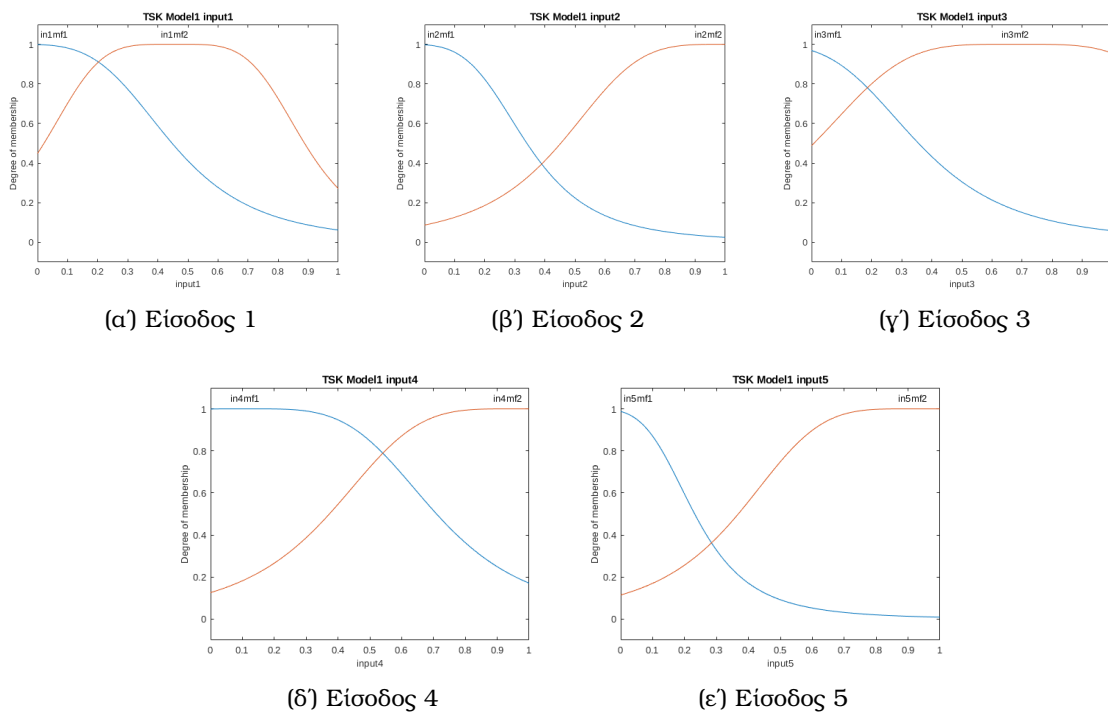
Αρχικοποιούμε τα μοντέλα βάσει των προδιαγραφών του Πίνακα 1.1, και ύστερα από την εκπαίδευσή τους, για το καθένα παρουσιάζουμε τα εξής ζητούμενα:

- Τις τελικές μορφές των ασαφών συνόλων που προέκυψαν μέσω της διαδικασίας εκπαίδευσης.
- Τα διαγράμματα μάθησης (Learning Curves του κάθε μοντέλου.
- Διάγραμμα σφαλμάτων πρόβλεψης.
- Τέλος, για όλα τα μοντέλα, οι τιμές των δεικτών απόδοσης *RMSE*, *NSME*, *NDEI*, R^2

Κάθε μοντέλο εκπαιδεύεται για 100 Epochs.

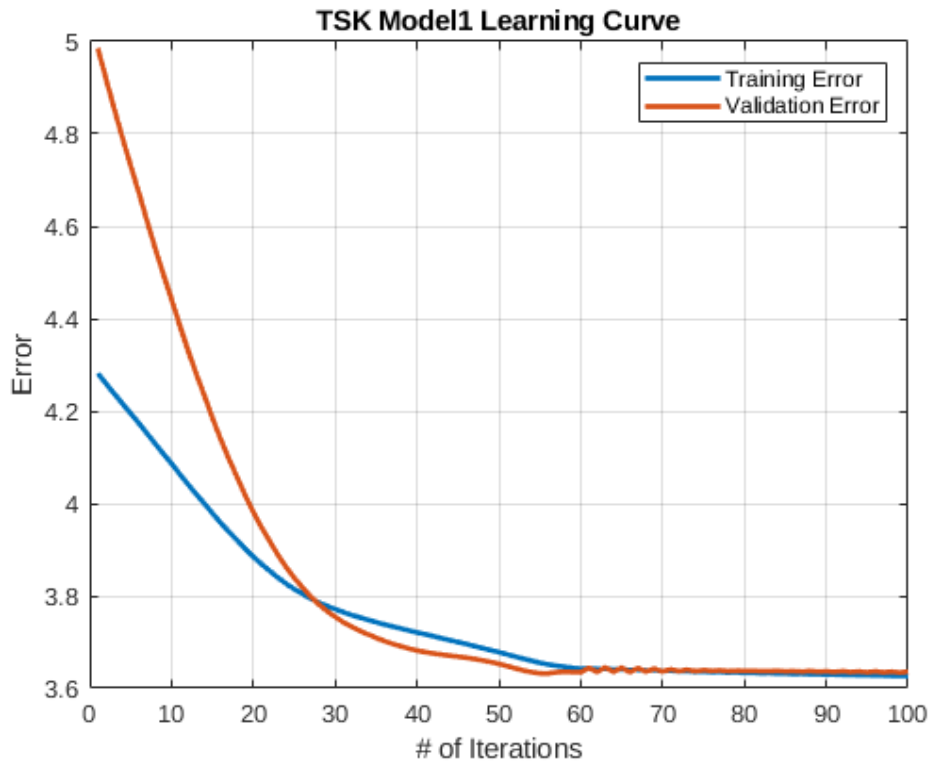
1.3.1 Μοντέλο 1 (2 MFs, Singleton)

- Τελικές Μορφές Ασαφών Συνόλων



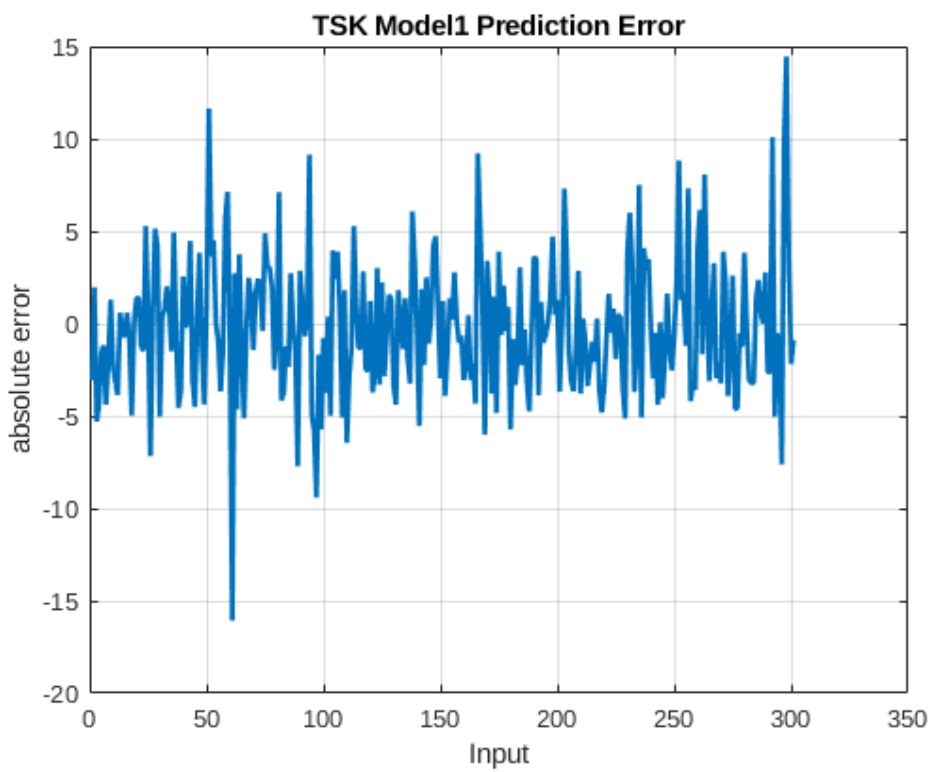
Σχήμα 1.2: Τελική Μορφή Ασαφών Συνόλων Μοντέλου 1

- Διάγραμμα μάθησης (Learning Curve)



Σχήμα 1.3: Διάγραμμα Μάθησης Μοντέλου 1

- Σφάλματα Πρόβλεψης

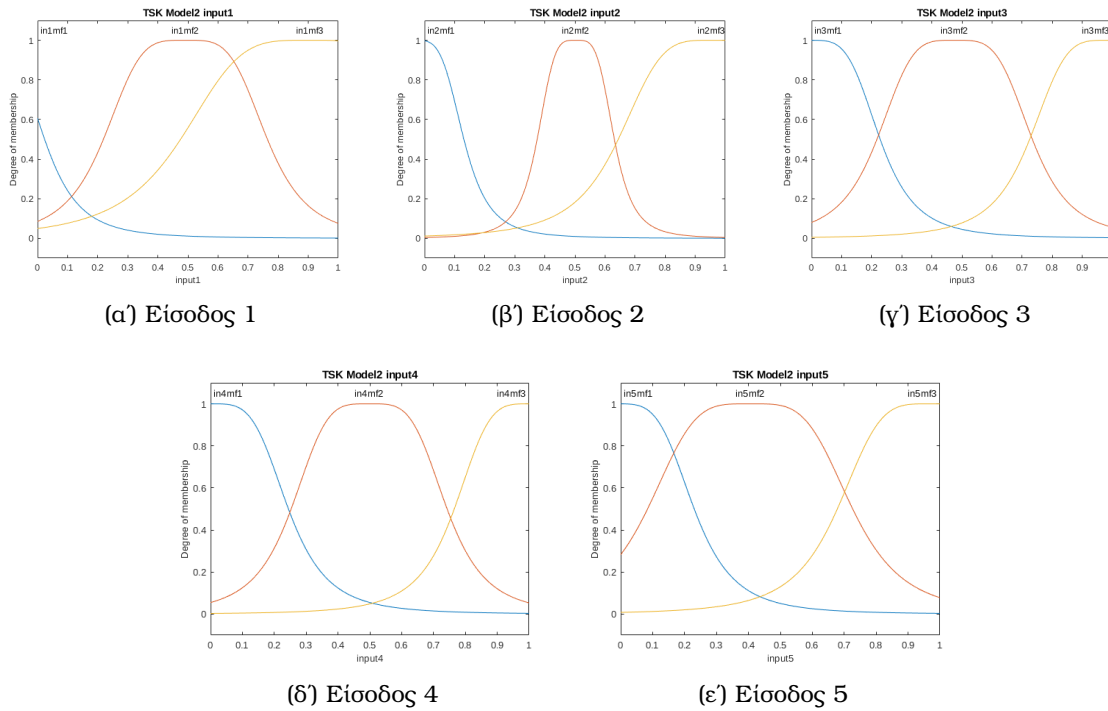


Σχήμα 1.4: Σφάλματα Πρόβλεψης Μοντέλου 1

Σύμφωνα με το Διάγραμμα Μάθησης, παρατηρούμε πώς μετά από έναν αριθμό επαναλήψεων, το Validation Error σε σχέση με το Training Error αρχίζει να συγκλίνει αρκετά. Το γεγονός ότι είναι πολύ κοντά τα σφάλματα μεταξύ τους είναι σημάδι ότι ο διαχωρισμός των δεδομένων ήταν αποτελεσματικός.

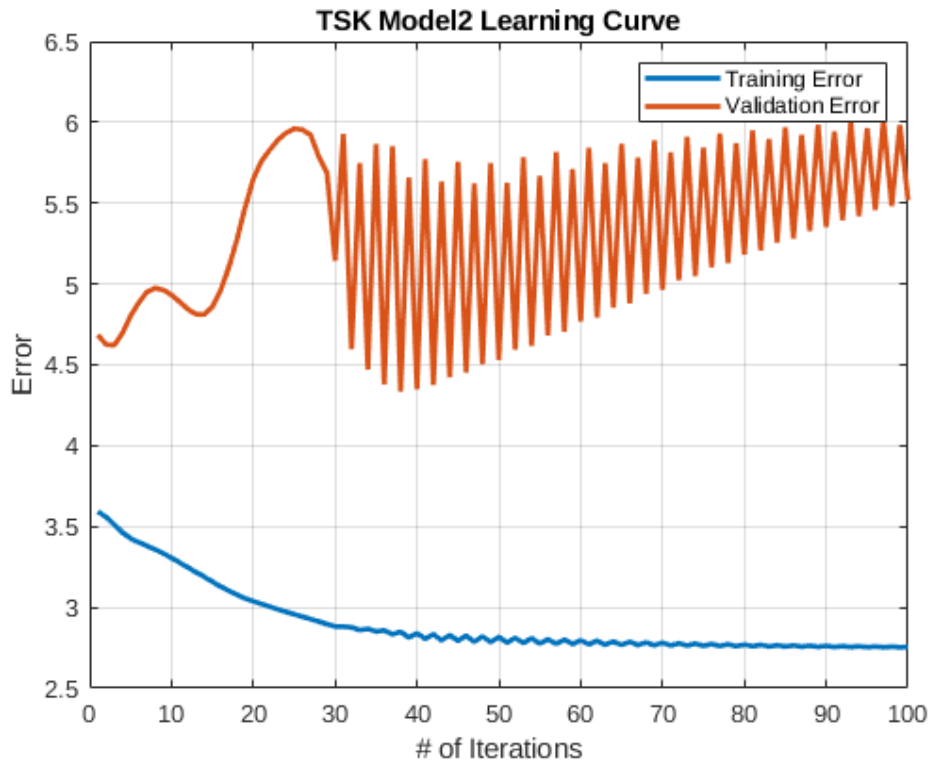
1.3.2 Μοντέλο 2 (3 MFs, Singleton)

- Τελικές Μορφές Ασαφών Συνόλων



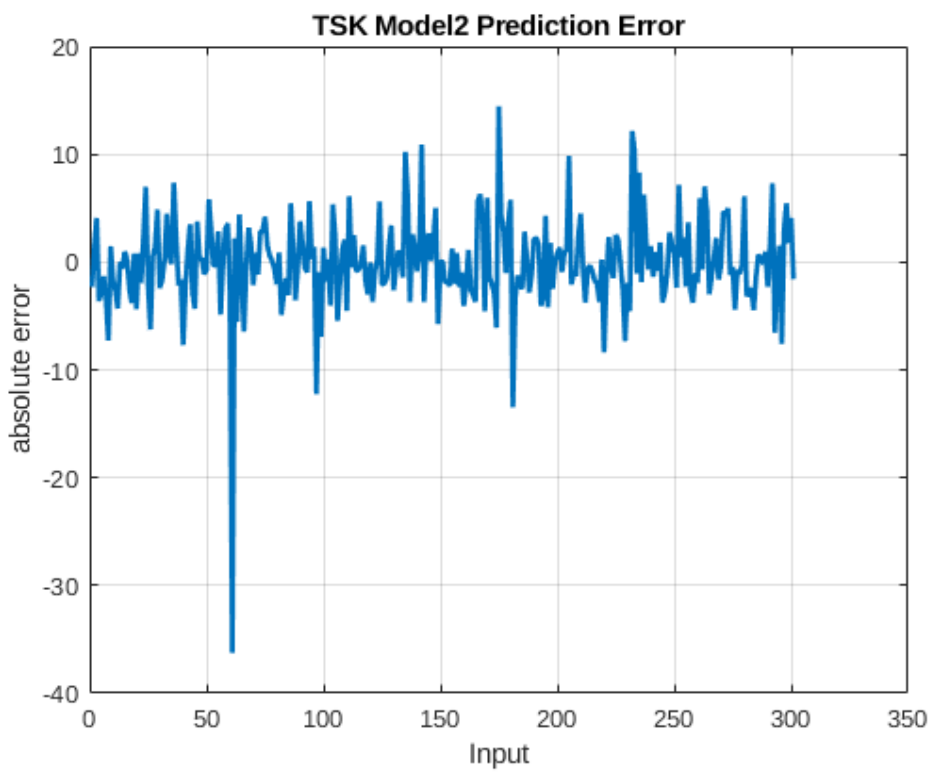
Σχήμα 1.5: Τελική Μορφή Ασαφών Συνόλων Μοντέλου 2

- Διάγραμμα μάθησης (Learning Curve)



Σχήμα 1.6: Διάγραμμα Μάθησης Μοντέλου 2

- Σφάλματα Πρόβλεψης

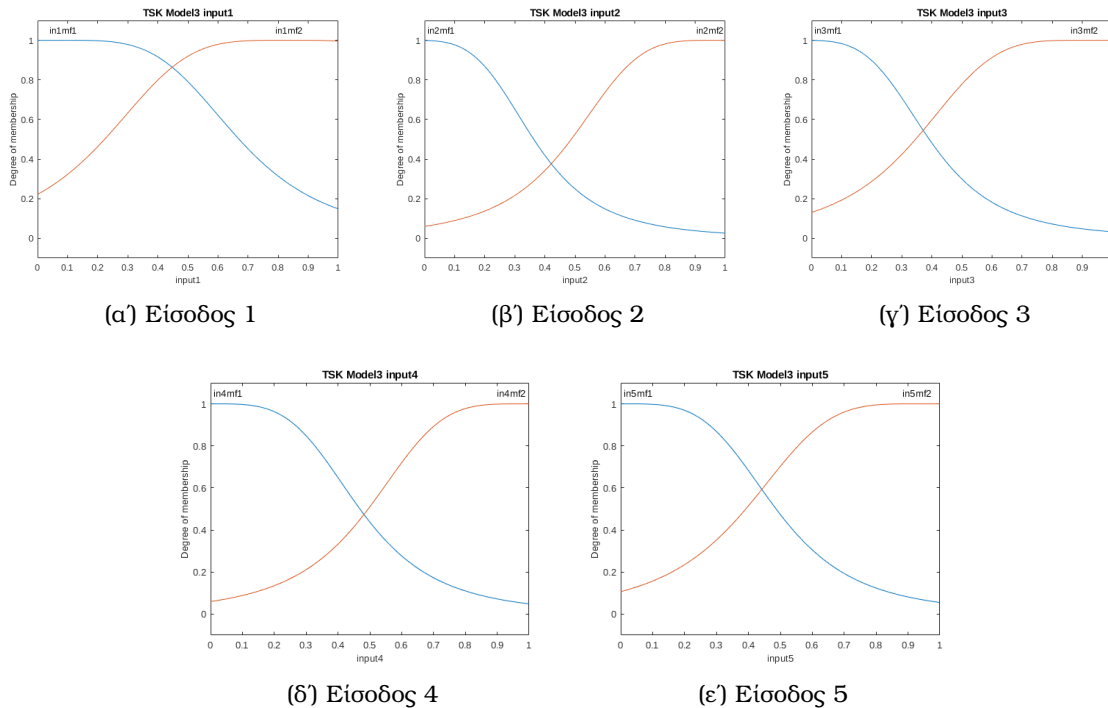


Σχήμα 1.7: Σφάλματα Πρόβλεψης Μοντέλου 2

Από το διάγραμμα μάθησης, του δεύτερου μοντέλου, παρατηρούμε, παρά τις ταλαντώσεις, ότι η απόκλιση του validation error σε σχέση με το training error αυξάνεται. Αυτό το γεγονός οφείλεται στο ότι στο μοντέλο εμφανίζεται φαινόμενο υπερεκπαίδευσης (overfitting).

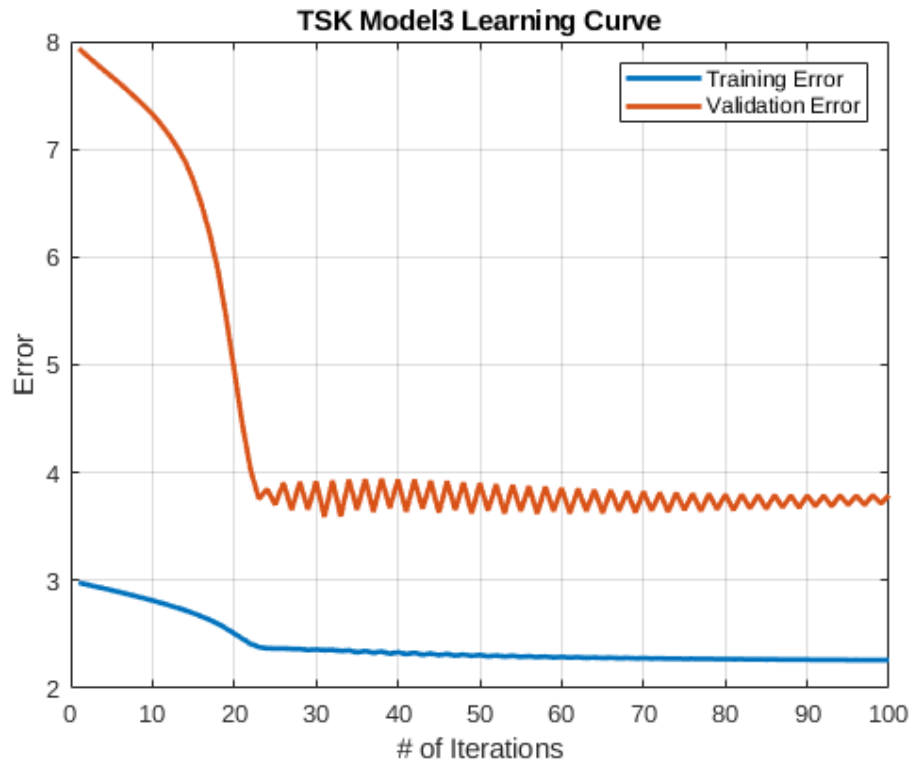
1.3.3 Μοντέλο 3 (2 MFs, Polynomial)

- Τελικές Μορφές Ασαφών Συνόλων



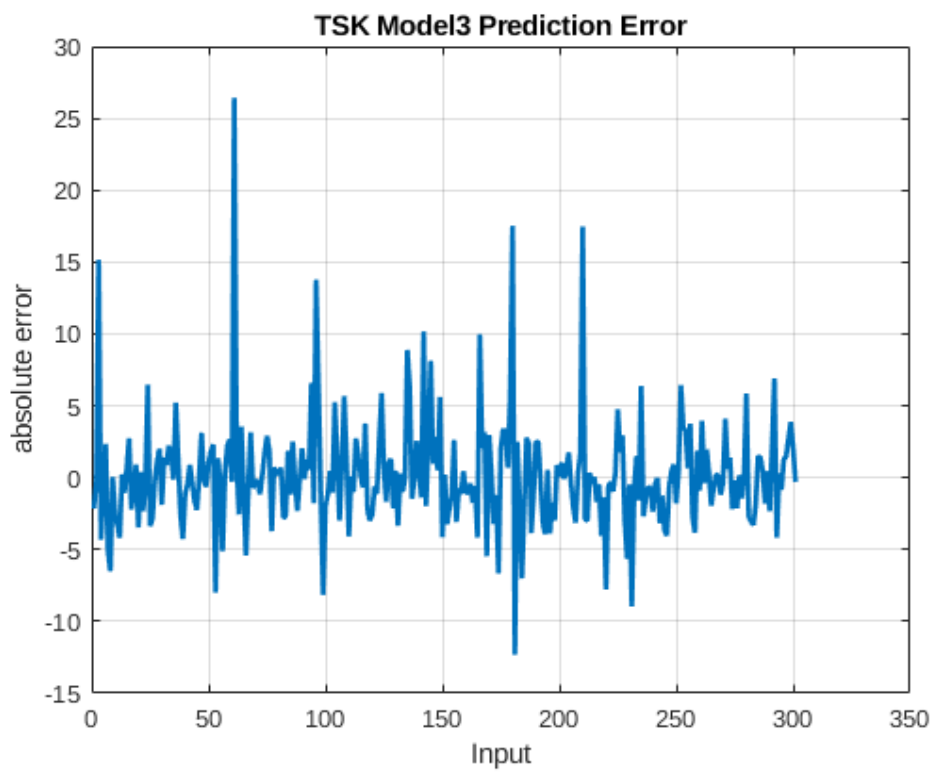
Σχήμα 1.8: Τελική Μορφή Ασαφών Συνόλων Μοντέλου 3

- Διάγραμμα μάθησης (Learning Curve)



Σχήμα 1.9: Διάγραμμα Μάθησης Μοντέλου 3

- Σφάλματα Πρόβλεψης

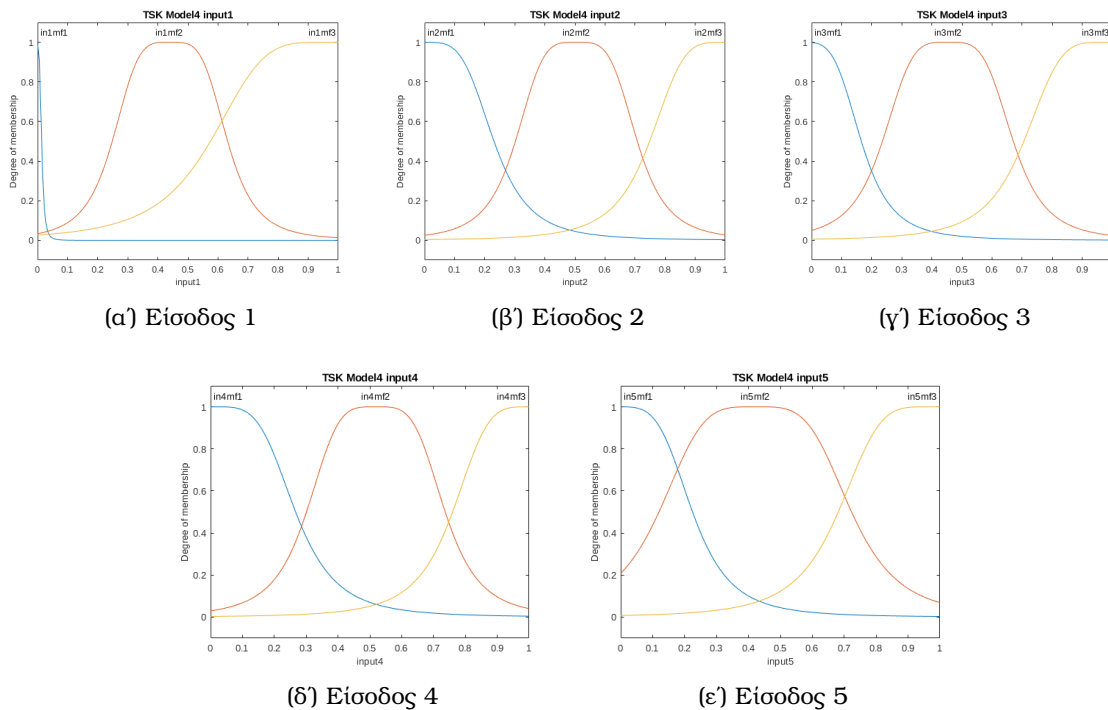


Σχήμα 1.10: Σφάλματα Πρόβλεψης Μοντέλου 3

Σε σύγκριση με το Μοντέλο 1, που έχει τα ίδια χαρακτηριστικά με μόνη διαφορά τον τύπο της μεταβλητής εξόδου, παρατηρούμε πως το Validation Error συγκλίνει πιο γρήγορα προς μια σταθερή τιμή. Παρ' όλα αυτά η απόκλιση μεταξύ των σφαλμάτων είναι αρκετή και παρατηρούμε και στο διάγραμμα με τα σφάλματα πρόβλεψης πως η διακύμανση αυτών είναι μεγαλύτερη.

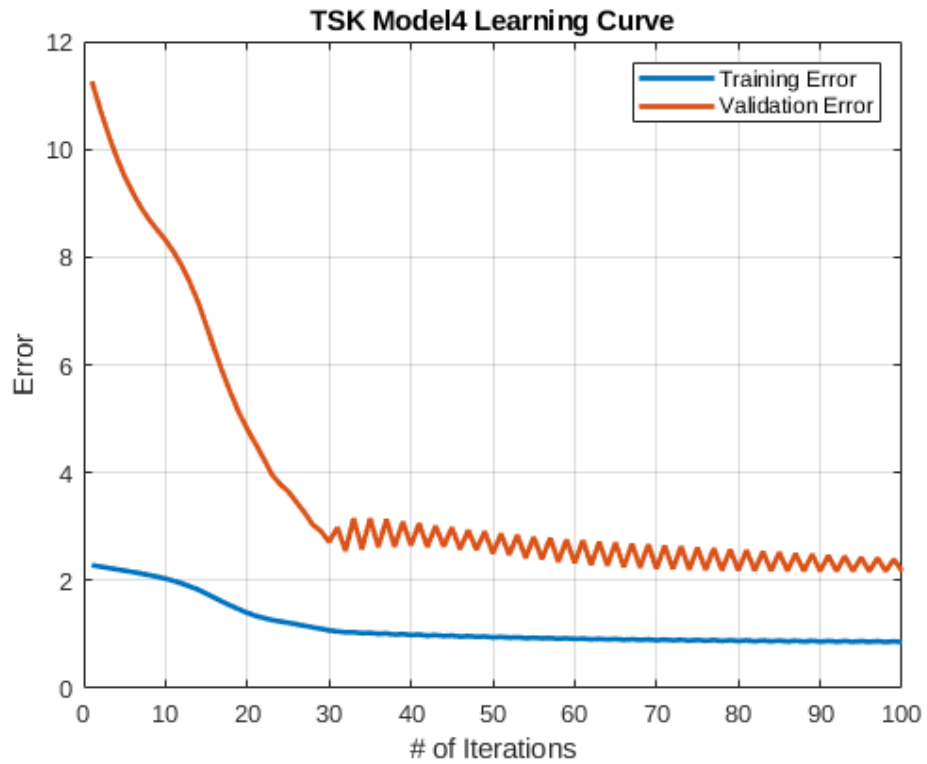
1.3.4 Μοντέλο 4 (3 MFs, Polynomial)

- Τελικές Μορφές Ασαφών Συνόλων



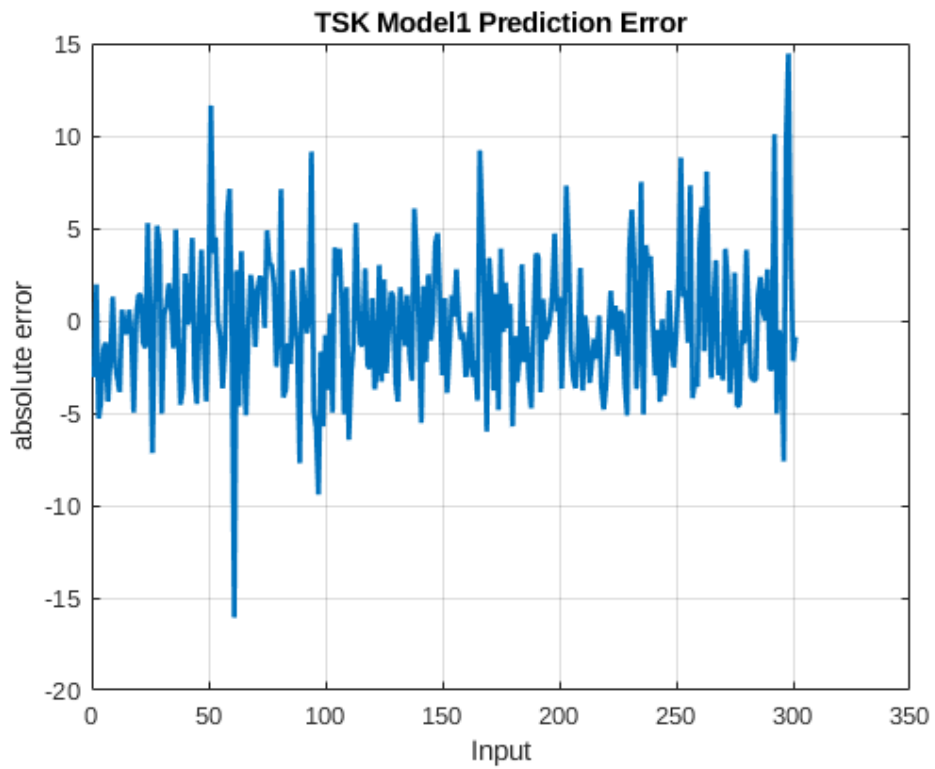
Σχήμα 1.11: Τελική Μορφή Ασαφών Συνόλων Μοντέλου 4

- Διάγραμμα μάθησης (Learning Curve)



Σχήμα 1.12: Διάγραμμα Μάθησης Μοντέλου 4

- Σφάλματα Πρόβλεψης



Σχήμα 1.13: Σφάλματα Πρόβλεψης Μοντέλου 4

Από το διάγραμμα μάθησης, βλέπουμε πως μετά την 30η επανάληψη το μοντέλο σταματάει να συγκλίνει, επομένως δεν είναι απαραίτητες οι υπόλοιπες επαναλήψεις εκμάθησης του μοντέλου. Επίσης η σύγκλιση του Validation Error σε σχέση με το Training Error είναι μικρότερη από το μοντέλο 3.

1.4 Αξιολόγηση

Στον παρακάτω πίνακα βλέπουμε τα αποτελέσματα των μετρικών Αξιολόγησης των μοντέλων:

Μετρικές Αξιολόγησης Μοντέλων				
	RMSE	NMSE	NDEI	R^2
TSK_model_1	3.6564	0.26757	0.51727	0.73243
TSK_model_2	4.2065	0.35413	0.59509	0.64587
TSK_model_3	3.8161	0.29146	0.53987	0.70854
TSK_model_4	2.6471	0.14024	0.37448	0.85976

Πίνακας 1.2: Πίνακας Αξιολόγησης

Τα αποτελέσματα σε bold, υποδεικνύουν την καλύτερη περίπτωση για κάθε μετρική. Βλέπουμε πως το TSK_Mdel_4 δείχνει να παρουσιάζει τις μικρότερες τιμές σφάλματος σε σύγκριση με τα υπόλοιπα μοντέλα. Σύμφωνα με τα παραπάνω, παρατηρούμε πως η συμπεριφορά των μοντέλων με έξοδο Polynomial είναι σχετικά καλύτερη από τα μοντέλα με Singleton έξοδο, δεδομένου ότι στο Μοντέλο 1, τα σφάλματα εκμάθησης και επικύρωσης ήταν αισθητά μεγαλύτερα από τα υπόλοιπα μοντέλα (υποψίες για underfitting). Όμοια, παρατηρείται πιο ικανοποιητική συμπεριφορά των μοντέλων για εισόδους με 3 συναρτήσεις συμμετοχής, έναντι 2. Βέβαια η αύξηση της απόδοσης του μοντέλου μπορεί να θεωρηθεί και αμελητέα, σε σύγκριση με την ραγδαία αύξηση των κανόνων IF . . . THEN του μοντέλου (από 32 σε 243!).

Κεφάλαιο 2

Εφαρμογή σε δεδομένα υψηλής διαστασιμότητας

Ένα προφανές πρόβλημα που ανακύπτει από την επιλογή Dataset με υψηλό βαθμό διαστασιμότητας είναι η λεγόμενη “έκρηξη” του πλήθους των IF-THEN κανόνων (*rule explosion*) όπως παρατηρήσαμε και προηγουμένως. Όπως είναι γνωστό από τη θεωρία, για την κλασσική περίπτωση του Grid Partitioning του χώρου εισόδου, ο αριθμός των κανόνων αυξάνεται εκθετικά σ σχέση με το πλήθος των εισόδων, γεγονός που καθιστά πολύ δύσκολη την μοντελοποίηση μέσω ενός TSK μοντέλου ακόμα και για datasets μεσαίας κλίμακας.

2.1 Προδιαγραφές

Το dataset που επιλέγεται για την εργασία είναι το Superconductivity dataset από το UCI Repository, το οποίο περιλαμβάνει 21.263 δείγματα που το καθένα περιγράφεται από 81 μεταβλητές/χαρακτηριστικά.

Για να αποφύγουμε το φαινόμενο rule explosion, καταφεύγουμε στην επιλογή χαρακτηριστικών και στην χρήση διαμέρισης διασκορπισμού. Οι δύο αυτές μέθοδοι εισάγουν στο πρόβλημα δύο ελεύθερες μεταβλητές:

- Τον αριθμό χαρακτηριστικών που θα χρησιμοποιήσουμε για την εκπαίδευση των μοντέλων.
- Την ακτίνα των clusters που καθορίζει την ακτίνα επιρροής των clusters και κατ’ επέκταση το πλήθος των κανόνων που θα προκύψουν.

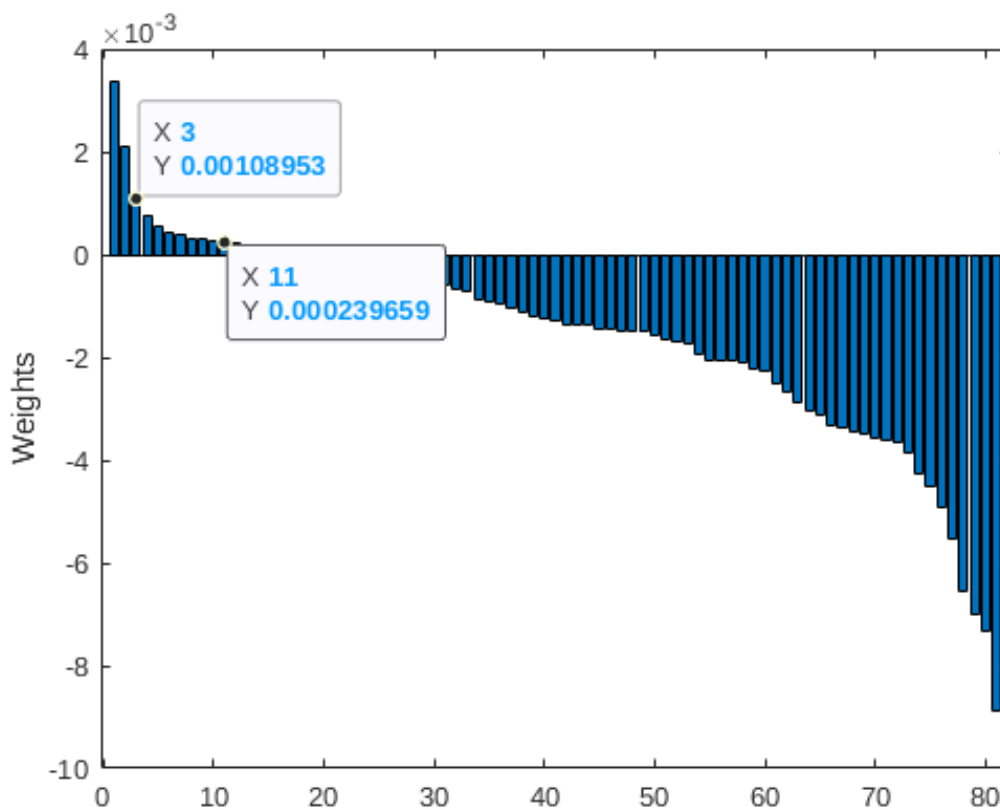
Για την επιλογή των βέλτιστων παραμέτρων ακολουθούμε την μέθοδο του grid search, δηλαδή δημιουργούμε ένα n -διάστατο πλέγμα (στην προκειμένη περίπτωση δισδιάστατο) όπου κάθε σημείο του πλέγματος αντιστοιχεί σε μια n -άδα τιμών για τις εν λόγω παραμέτρους. Σε αυτά τα σημεία χρησιμοποιούμε μια μέθοδο αξιολόγησης για να ελέγξουμε την ορθότητα των συγκεκριμένων τιμών.

Η μέθοδος αξιολόγησης που χρησιμοποιούμε είναι η διασταυρωμένη επικύρωση. Δηλαδή, για κάθε επιλεγμένη τιμή παραμέτρων, χωρίζουμε το σύνολο εκπαίδευσης σε δύο υποσύνολα, από τα οποία το ένα θα χρησιμοποιηθεί για την εκπαίδευση ενός μοντέλου και το άλλο για την αξιολόγησή του. Η διαδικασία αυτή επαναλαμβάνεται για

διαφορετικό διαχωρισμό του συνόλου κάθε φορά (συνήθως 5 ή 10 φορές) όπου τελικά λαμβάνουμε το μέσο όρο του σφάλματος του μοντέλου. Για κάθε διαδικασία διασταυρωμένης επικύρωσης, τα μοντέλα προπονούνται για 75 Epochs.

2.2 Grid Search

Για την επιλογή χαρακτηριστικών, εφαρμόζουμε στα δεδομένα μας την συνάρτηση του MATLAB, *relieff* για *k*-nearest neighbors. Η συνάρτηση αυτή επιστρέφει τους δείκτες των χαρακτηριστικών των δειγμάτων με σειρά σημαντικότητας βάσει ενός βάρους, όπως φαίνεται στο Σχ.2.1. Παρατηρούμε πως ήδη μέχρι το 11ο σημαντικότερο χαρακτηριστικό, το βάρος είναι πολύ μικρό. Επομένως δεν θα χρειαστεί να χρησιμοποιήσουμε περισσότερα χαρακτηριστικά. Επίσης για λόγους πρακτικότητας δεν ξεπερνάμε αυτήν την τιμή διότι ο χρόνος εκπαίδευσης αυξάνεται εκθετικά στο περιβάλλον MATLAB, σε σχέση με τον αριθμό χαρακτηριστικών. Επομένως για τις παραμέτρους του πλέγματος



Σχήμα 2.1: Βάρη σημαντικότητας χαρακτηριστικών

διαλέγουμε:

- Αριθμός χαρακτηριστικών = {5, 7, 9, 11},
- ακτίνα cluster $r_{cluster} = \{0.3, 0.6, 0.9\}$

Η επιλογή των ακτινών έγινε αυθαίρετα με την λογική ότι για κάθε αριθμό επιλεγμένων χαρακτηριστικών, θα μελετήσουμε ένα μοντέλο με μικρή, μεσαία και μεγάλη τιμή

RMSE	$feat_{num} = 5$	$feat_{num} = 7$	$feat_{num} = 9$	$feat_{num} = 11$
$r_{cluster} = 0.3$	16.4308	15.4119	16.2610	15.0812
$r_{cluster} = 0.6$	18.8659	17.1757	17.8375	16.6174
$r_{cluster} = 0.9$	20.3239	19.0822	18.6452	18.0366

(α) RMSE Matrix

NMSE	$feat_{num} = 5$	$feat_{num} = 7$	$feat_{num} = 9$	$feat_{num} = 11$
$r_{cluster} = 0.3$	0.2332	0.2055	0.2302	0.1965
$r_{cluster} = 0.6$	0.3074	0.2549	0.2750	0.2386
$r_{cluster} = 0.9$	0.3568	0.3145	0.3003	0.2810

(β) NMSE Matrix

NDEI	$feat_{num} = 5$	$feat_{num} = 7$	$feat_{num} = 9$	$feat_{num} = 11$
$r_{cluster} = 0.3$	0.4829	0.4529	0.4779	0.4432
$r_{cluster} = 0.6$	0.5545	0.5048	0.5242	0.4884
$r_{cluster} = 0.9$	0.5973	0.5608	0.5480	0.5301

(γ) NDEI Matrix

R^2	$feat_{num} = 5$	$feat_{num} = 7$	$feat_{num} = 9$	$feat_{num} = 11$
$r_{cluster} = 0.3$	0.7668	0.7945	0.7698	0.8035
$r_{cluster} = 0.6$	0.6926	0.7451	0.7250	0.7614
$r_{cluster} = 0.9$	0.6432	0.6855	0.6997	0.7190

(δ) R^2 Matrix

Πίνακας 2.1: Πίνακες Δεικτών Μοντέλων Πλέγματος

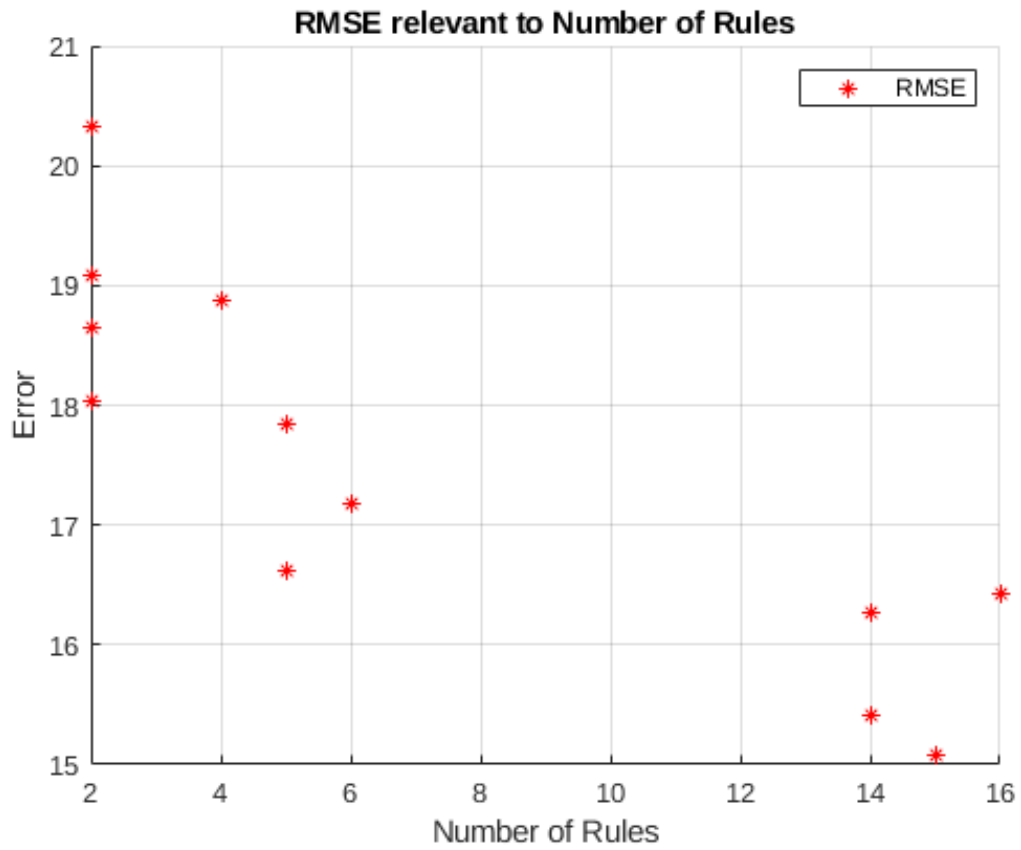
ακτίνας $r_{cluster}$

Υστερα από την εκτέλεση της αναζήτησης πλέγματος και αξιολόγησης μέσω 5-πτυχς διασταυρωμένης επικύρωσης (*5-fold cross validation*) υπολογίζουμε τους δείκτες απόδοσης των μοντέλων των πινάκων 2.1:

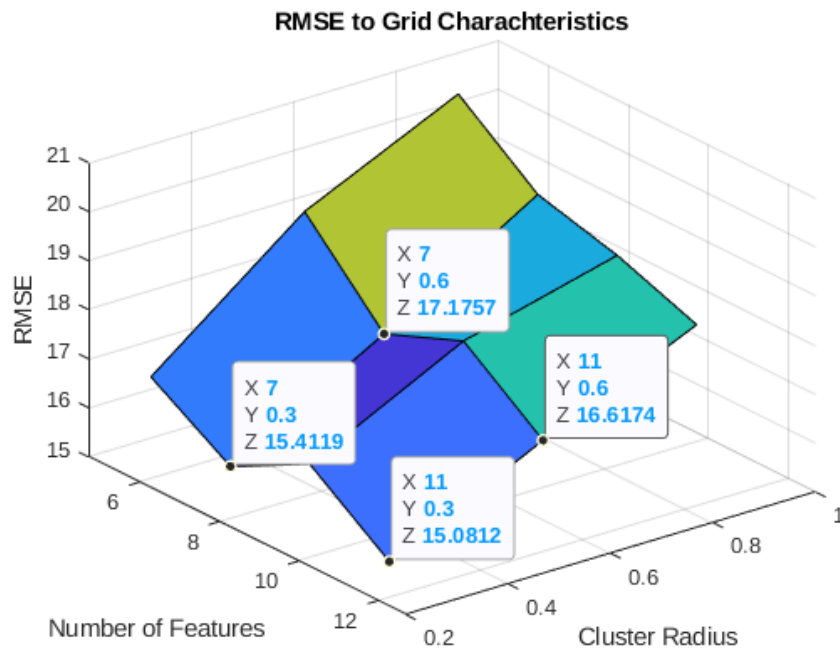
Στο Σχ.2.2 βλέπουμε την σχέση της ρίζας του μέσου τετραγωνικού σφάλματος του μοντέλου με το πλήθος των κανόνων που παράγει το κάθε μοντέλο. Στο σχήμα έχουμε 12 μετρήσεις καθώς το σφάλμα που μετράμε είναι το μέσο σφάλμα από κάθε k-fold. Παρατηρούμε πως καθώς αυξάνεται το πλήθος των κανόνων, μειώνεται το μέσο σφάλμα του μοντέλου. Διαισθητικά παρατηρούμε ένα trade-off μεταξύ της ερμηνευσιμότητας (interpretability) και της ακρίβειας (accuracy) του μοντέλου. Δηλαδή μειώνοντας την ευκολία ερμηνείας του μοντέλου (αυξάνοντας την πολυπλοκότητα) παρατηρείται αντίστοιχη αύξηση στην ακρίβεια του μοντέλου, και αντίστροφα.

2.3 Βέλτιστο μοντέλο

Στον Πίνακα 2.1 σημειώνουμε τις βέλτιστες τιμές για κάθε δείκτη με bold. Για κάθε μετρική βλέπουμε πως η βέλτιστη τιμή ανήκει στο σημείο πλέγματος με αριθμό χαρακτηριστικών = 11 και με ακτίνα cluster = 0.3. Αυτό μπορούμε να το διαπιστώσουμε και από την επιφάνεια του Σχ.2.3, στην οποία φαίνονται τα υποψήφια σημεία εύρεσης του βέλτιστου μοντέλου.



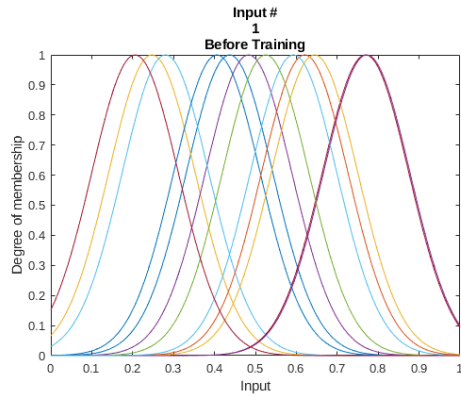
Σχήμα 2.2: RMSE συναρτήσει του πλήθους κανόνων



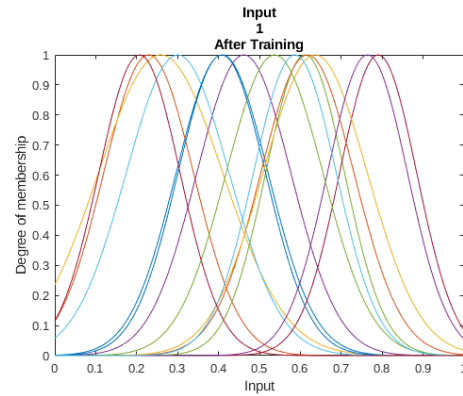
Σχήμα 2.3: Επιφάνεια Σφάλματος συναρτήσει παραμέτρων αναζήτησης πλέγματος

Εφόσον διαλέγουμε μοντέλο με αριθμό χαρακτηριστικών $feat_{num} = 11$ και μέγεθος ακτίνας cluster, $r_{cluster} = 0.3$, το εκπαιδεύουμε για τα αρχικά δείγματα εκπαίδευσης αλλά κρατάμε μόνο τα 11 πιο σημαντικά χαρακτηριστικά. Από την εκπαίδευση του μοντέλου έχουμε:

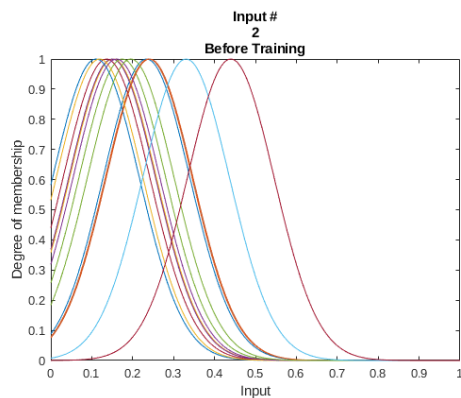
- Αρχική και τελική μορφή μερικών Ασαφών Συνόλων



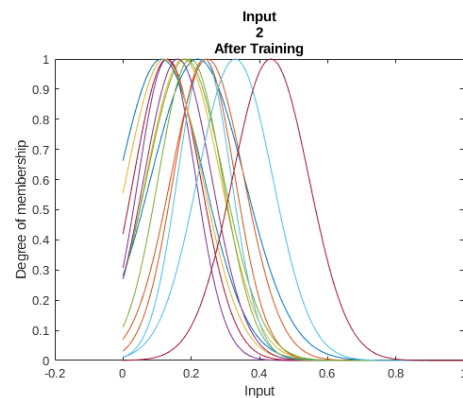
(α) Είσοδος 1 αρχικό



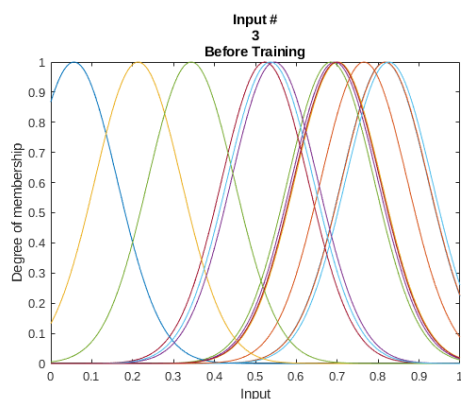
(β) Είσοδος 1 τελικό



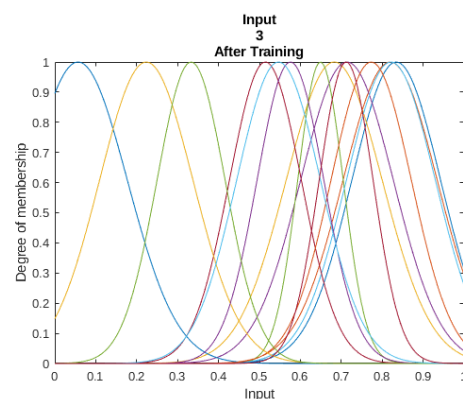
(γ) Είσοδος 2 αρχικό



(δ) Είσοδος 2 τελικό



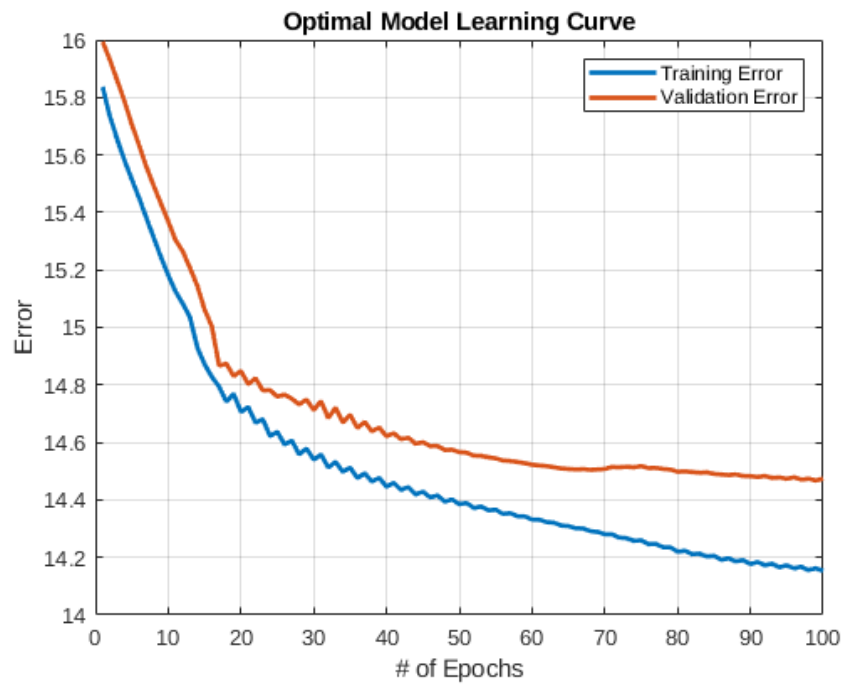
(ε) Είσοδος 3 αρχικό



(ς) Είσοδος 3 τελικό

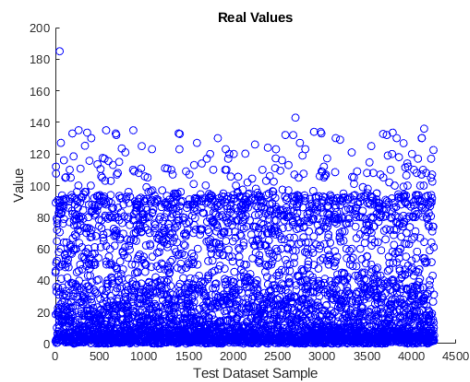
Σχήμα 2.4: Αρχικές και τελικές μορφές Ασαφών Συνόλων

- Διάγραμμα Μάθησης

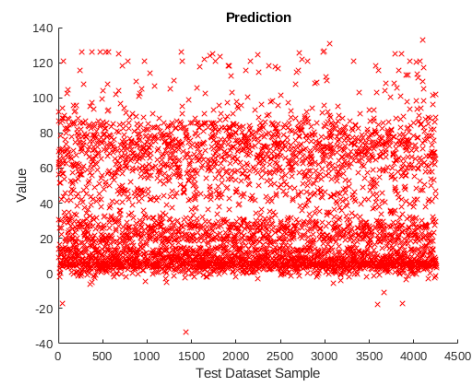


Σχήμα 2.5: Καμπύλη Εκπαίδευσης Βέλτιστου Μοντέλου

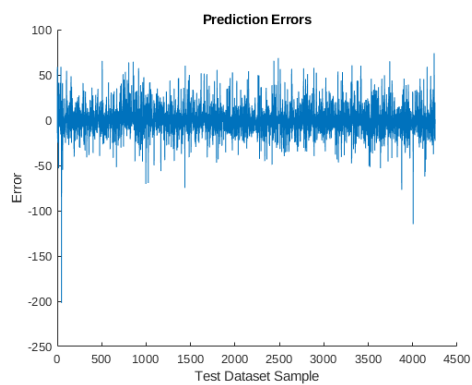
- Διαγράμματα Πραγματικών/Προβλεπόμενων τιμών



(α') Πραγματικές τιμές



(β') Προβλέψεις Μοντέλου



(γ') Σφάλμα Πρόβλεψης

Σχήμα 2.6: Πραγματικές τιμές / Προβλέψεις μοντέλου

- Τιμές δεικτών απόδοσης

Δείκτες Βέλτιστου Μοντέλου			
RMSE	NMSE	NDEI	R²
14.921	0.19231	0.43853	0.80769

Πίνακας 2.2: Δείκτες Απόδοσης Βέλτιστου Μοντέλου

Βλέπουμε από τα διαγράμματα σφάλματος πρόβλεψης πως το μοντέλο, προβλέπει σε ικανοποιητικό βαθμό το σύνολο ελέγχου, με μεμονωμένες περιπτώσεις μεγάλων σφαλμάτων. Το σύνολο των κανόνων του μοντέλου είναι 15 κανόνες. Για τον ίδιο αριθμό κρατημένων χαρακτηριστικών με την μέθοδο του grid partitioning, αν για κάθε είσοδο είχαμε 2 ή 3 ασαφή σύνολα, θα είχαμε συνολικά 2^{15} ή 3^{15} κανόνες αντίστοιχα. Αυτή η τεράστια διαφορά μας επιβεβαιώνουν πόσο μη πρακτικό και χρονοβόρο στη φάση της εκπαίδευσης, θα ήταν το μοντέλο μας, αν αντί για την μέθοδο του subtractive clustering χρησιμοποιούσαμε τη μέθοδο του grid partitioning.

Παραδοτέο

Περιεχόμενα

Το παραδοτέο της εργασίας αποτελείται από:

- τρία *.m scripts*
 - ***TSKRegression_simple.m***, στο οποίο είναι γραμμένη η υλοποίηση του πρώτου μέρους της εργασίας,
 - ***TSKRegression_dimensionality.m***, στο οποίο υλοποιείται το δεύτερο μέρος της εργασίας,
 - ***split_scale.m***, βοηθητική συνάρτηση διαχώρισης των αρχικών dataset, του elearning.
- δύο *.mat scripts*
 - ***simpleReg.mat*** στο οποίο είναι αποθηκευμένες οι μεταβλητές του Workspace που παρήχθησαν από το *TSKRegression_simple.m* και χρησιμοποιήθηκαν για την συγγραφή του πρώτου μέρους της εργασίας.
 - ***dimReg.mat***, στο οποίο είναι αποθηκευμένες οι μεταβλητές του Workspace που παρήχθησαν από το *TSKRegression_dimensionality.m* και χρησιμοποιήθηκαν για την συγγραφή του δεύτερου μέρους της εργασίας.
- την παρούσα αναφορά.

Περισσότερο υλικό, όπως η εκφώνηση της άσκησης, το αποθετήριο εικόνων και ο κώδικας σε L^AT_EX, βρίσκεται αποθηκευμένο στο [git-hub](#).

Τεχνικές Οδηγίες

Για την επιθεώρηση της εργασίας, εκτελείται αρχικά το αρχείο ***TSKRegression_simple.m*** και ύστερα το ***TSKRegression_dimensionality.m***. Διότι η διαδικασία εκτέλεση των αρχείων είναι χρονοβόρα, για διευκόλυνση των ελέγχων, έχουμε αποθηκεύσει τις μεταβλητές του Workspace που λάβαμε κατά την εκτέλεση των αλγορίθμων.

Επειδή δεν υπάρχει σταθερός διαμερισμός του αρχικού dataset, τα αποτελέσματα μπορεί να αποκλίνουν για κάθε ξεχωριστή εκτέλεση των αλγορίθμων.