



Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης
Πολυτεχνική Σχολή
Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

ΥΠΟΛΟΓΙΣΤΙΚΗ ΝΟΗΜΟΣΥΝΗ

ΤΡΙΤΗ ΕΡΓΑΣΤΗΡΙΑΚΗ ΑΣΚΗΣΗ

ΕΠΙΛΥΣΗ ΠΡΟΒΛΗΜΑΤΟΣ ΠΑΛΙΝΔΡΟΜΗΣΗΣ ΜΕ ΧΡΗΣΗ ΜΟΝΤΕΛΩΝ TSK

Ημερομηνία: 11/10/2024

Οικονόμου Χρήστος

A.E.M.: 10268

Email: cn οικον ο μ @ ece. a u t h . g r

ΠΕΡΙΕΧΟΜΕΝΑ

Εισαγωγή.....	σελ.3
Μοντέλα TSK	3
Στόχος της Εργασίας	3
 Πρώτο Μέρος - Εφαρμογή σε Απλό Dataset	4
Μοντέλο 1 (TSK_Model_1)	5
Μοντέλο 2 (TSK_Model_2)	7
Μοντέλο 3 (TSK_Model_3)	9
Μοντέλο 4 (TSK_Model_4)	11
Τιμές Δεικτών Απόδοσης	14
Συμπεράσματα	14
 Δεύτερο Μέρος - Εφαρμογή σε Dataset με Υψηλή Διαστασιμότητα	15
Συμπεράσματα	21
 Βιβλιογραφικές Πηγές	21

Εισαγωγή

Μοντέλα TSK

Τα μοντέλα TSK (Takagi – Sugeno – Kang) αποτελούν έναν τύπο συστήματος, το οποίο παράγει ασαφείς (“fuzzy”) κανόνες με βάση ένα σετ δεδομένων που δίνεται ως είσοδος. Σχεδιάζονται για να προσεγγίζουν σύνθετα, μη γραμμικά συστήματα, χρησιμοποιώντας ένα σύνολο κανόνων “αν - τότε” (“if - then” rules), ενώ συνδυάζουν την ασαφή λογική με μαθηματικές συναρτήσεις για να δημιουργήσουν μια έξοδο, καθιστώντας τα δημοφιλή για συστήματα ελέγχου, πρόβλεψης και λήψης αποφάσεων.

Κύρια χαρακτηριστικά των μοντέλων TSK:

1. **Ασαφείς Κανόνες:** Όπως και άλλα συστήματα ασαφούς λογικής, τα μοντέλα TSK χρησιμοποιούν κανόνες “αν-τότε”. Το μέρος “αν” ορίζει ασαφή σύνολα για τις εισόδους, και το μέρος “τότε” ορίζει μια μαθηματική συνάρτηση (συνήθως γραμμική ή πολυωνυμική) που καθορίζει την έξοδο.
2. **Ακριβείς Έξοδοι:** Σε αντίθεση με τα παραδοσιακά συστήματα ασαφούς λογικής που παράγουν ασαφείς τιμές, τα μοντέλα TSK συνήθως παράγουν ακριβείς, πραγματικές τιμές. Αυτό επιτυγχάνεται με τον συνδυασμό των αποτελεσμάτων από όλους τους κανόνες.
3. **Δομή Κανόνων:** Η δομή των κανόνων ενός μοντέλου TSK είναι:
 - **Αν** (προηγούμενο): ασαφείς συνθήκες για τις μεταβλητές εισόδου.
 - **Τότε** (επόμενο): μια γραμμική ή πολυωνυμική συνάρτηση που εφαρμόζεται στις μεταβλητές εισόδου.
4. **Σταθμισμένος Μέσος για την Έξοδο:** Η συνολική έξοδος του συστήματος είναι ένας σταθμισμένος μέσος όρος των εξόδων όλων των κανόνων, με τα βάρη να καθορίζονται από τον βαθμό στον οποίο οι τιμές εισόδου ικανοποιούν τις ασαφείς συνθήκες στα προηγούμενα.

Στόχος της Εργασίας

Η παρούσα εργασία εξετάζει το κατά πόσο τα μοντέλα TSK έχουν την ικανότητα να μοντελοποιήσουν μη γραμμικές συναρτήσεις πολλών μεταβλητών, με σκοπό την επίλυση προβλημάτων παλινδρόμησης (regression). Για την εξαγωγή των συμπερασμάτων, γίνεται εφαρμογή των μοντέλων σε δύο datasets από το UCI Repository, με σκοπό την εκτίμηση της μεταβλητής στόχου από τα διαθέσιμα δεδομένα, με χρήση ασαφών νευρωνικών μοντέλων. Αναλυτικότερη εξήγηση της διαδικασίας παρατίθεται στις ενότητες που ακολουθούν.

Πρώτο Μέρος - Εφαρμογή σε Απλό Dataset

Στο πρώτο μέρος της εργασίας, εξετάζεται η εφαρμογή τεσσάρων μοντέλων TSK σε ένα σχετικά απλό σύνολο δεδομένων και γίνεται διερεύνηση του τρόπου εκπαίδευσης και αξιολόγησης των μοντέλων αυτών. Συγκεκριμένα, χρησιμοποιείται το σύνολο δεδομένων Airfoil Self-Noise από το UCI Repository, το οποίο περιλαμβάνει 1503 δείγματα (instances) και 6 χαρακτηριστικά (features). Ο κώδικας MATLAB που αποτελεί την υλοποίηση του μέρους αυτού της εργασίας είναι καταγεγραμμένος στο αρχείο “regression_a_10268.m”.

Αρχικά γίνεται ο κατάλληλος διαχωρισμός του dataset, με σκοπό τον σχηματισμό τριών υποσυνόλων δεδομένων: του υποσυνόλου **εκπαίδευσης**, του υποσυνόλου **επικύρωσης** και του υποσυνόλου **ελέγχου απόδοσης** του τελικού μοντέλου (D_{chk}). Ο διαχωρισμός έγινε μέσω της συνάρτησης `split_scale` (αρχείο “split_scale.m”), ως εξής:

- Υποσύνολο Εκπαίδευσης (D_{trn}) \rightarrow 60% του αρχικού dataset \rightarrow 902 instances
- Υποσύνολο Επικύρωσης (D_{val}) \rightarrow 20% του αρχικού dataset \rightarrow 300 instances
- Υποσύνολο Ελέγχου (D_{chk}) \rightarrow 20% του αρχικού dataset \rightarrow 301 instances

Επόμενο βήμα αποτέλεσε η **εκπαίδευση** των τεσσάρων TSK μοντέλων, στα οποία μεταβαλλόταν το πλήθος των συναρτήσεων συμμετοχής (Membership Functions – MFs), καθώς και η μορφή της εξόδου τους σύμφωνα με τον Πίνακα 1.

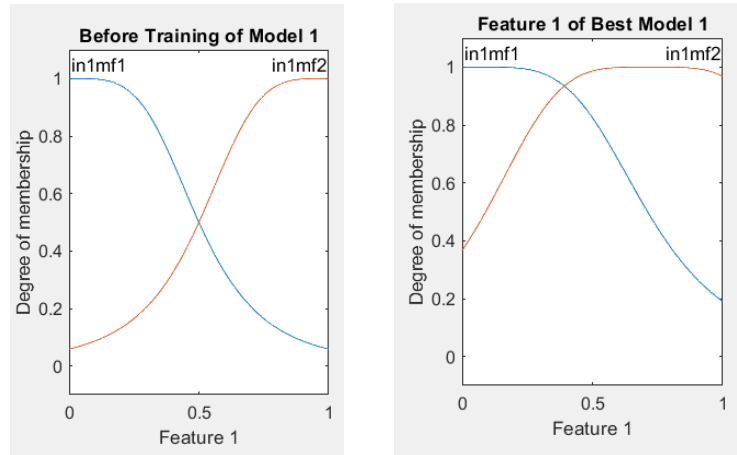
Πλήθος συναρτήσεων συμμετοχής		Μορφή εξόδου
TSK_model_1	2	Singleton
TSK_model_2	3	Singleton
TSK_model_3	2	Polynomial
TSK_model_4	3	Polynomial

Πίνακας 1: Ταξινόμηση Μοντέλων προς Εκπαίδευση

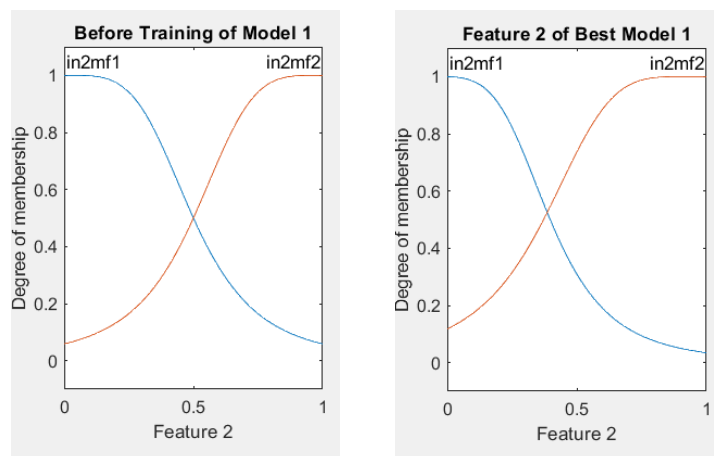
Τα μοντέλα εκπαιδεύτηκαν σύμφωνα με τα ζητούμενα και τις προδιαγραφές που προέβλεπε η εκφώνηση της εργασίας. Παρακάτω παρατίθενται τα σχετικά διαγράμματα για κάθε ένα από τα τέσσερα μοντέλα που εξετάζονται:

Μοντέλο 1 (TSK_model_1)

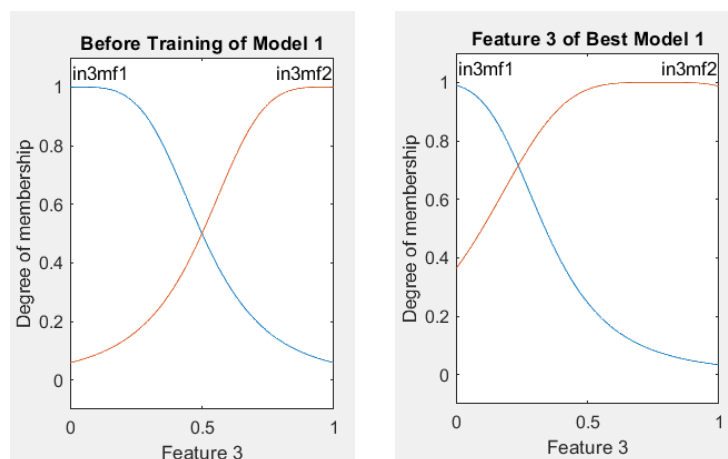
Διαγράμματα για κάθε Feature πριν και μετά την εκπαίδευση του Μοντέλου 1:



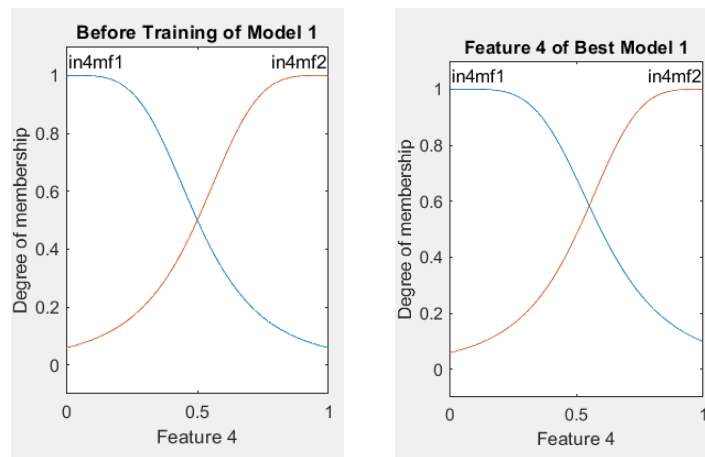
Διάγραμμα 1: TSK Model 1 – Feature 1



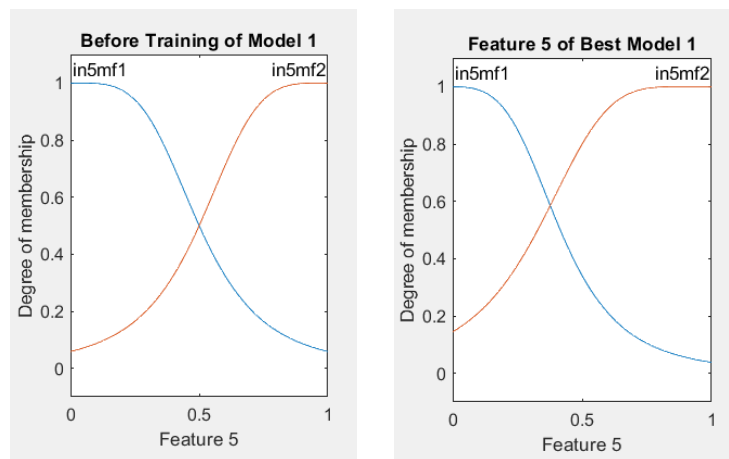
Διάγραμμα 2: TSK Model 1 – Feature 2



Διάγραμμα 3: TSK Model 1 – Feature 3

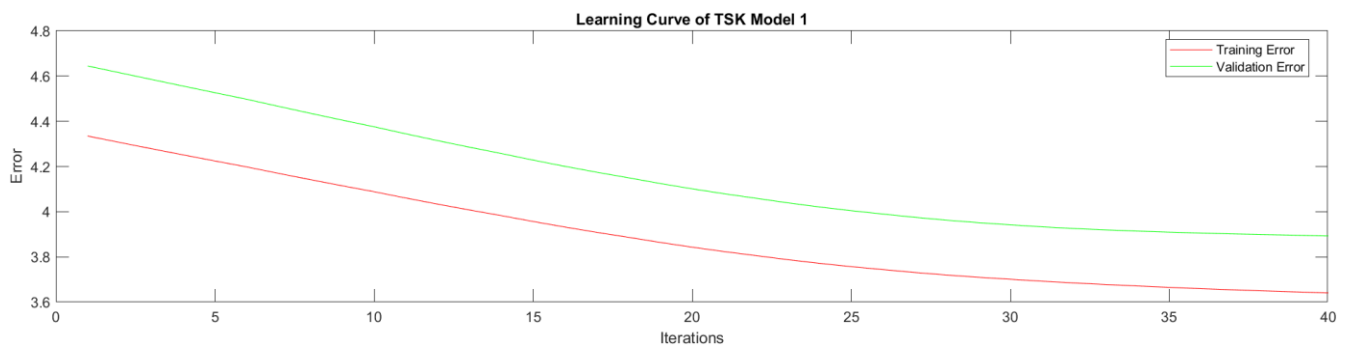


Διάγραμμα 4: TSK Model 1 – Feature 4

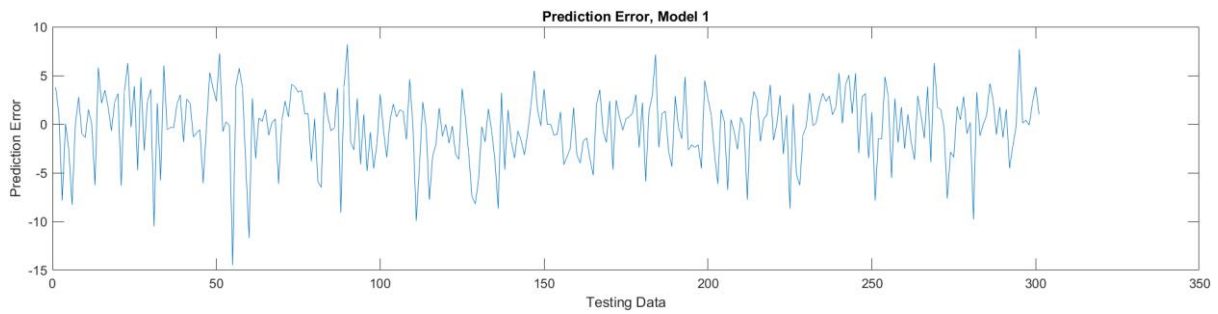


Διάγραμμα 5: TSK Model 1 – Feature 5

Καμπύλη Εκμάθησης και Σφάλμα Πρόβλεψης για το Μοντέλο 1:



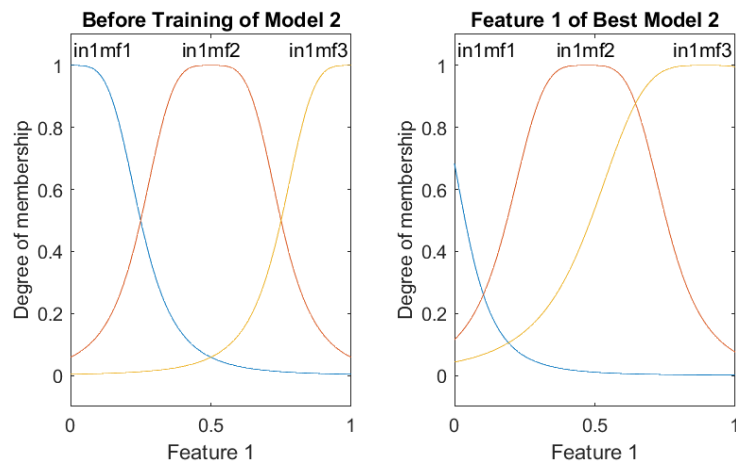
Διάγραμμα 6: TSK Model 1 – Καμπύλη Εκμάθησης



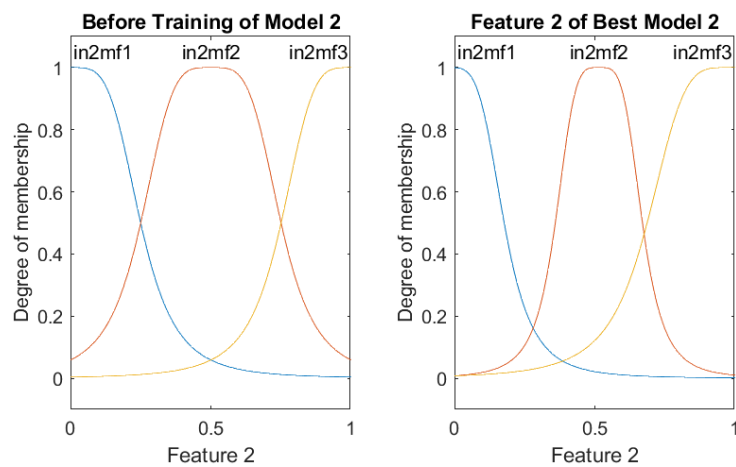
Διάγραμμα 7: TSK Model 1 – Σφάλμα Πρόβλεψης

Μοντέλο 2 (TSK_model_2)

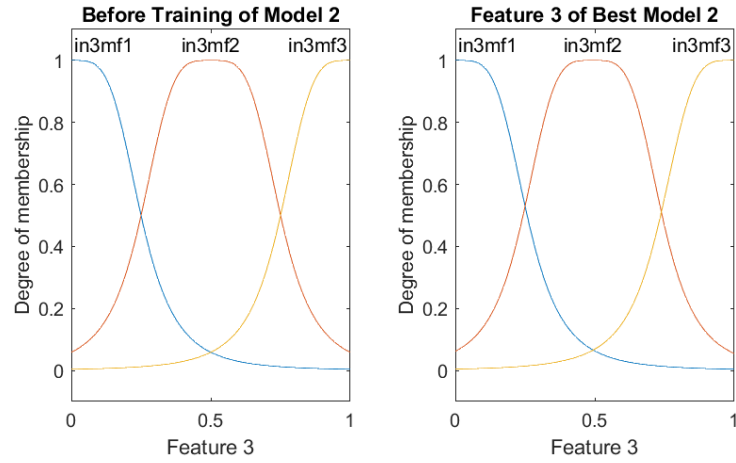
Διαγράμματα για κάθε Feature πριν και μετά την εκπαίδευση του Μοντέλου 2:



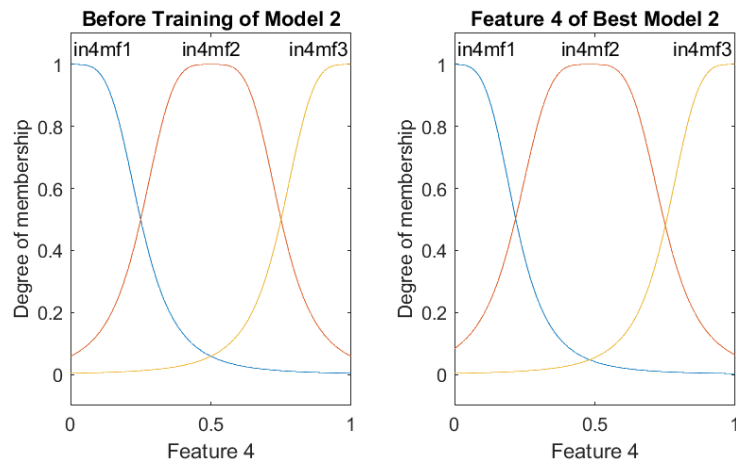
Διάγραμμα 8: TSK Model 2 – Feature 1



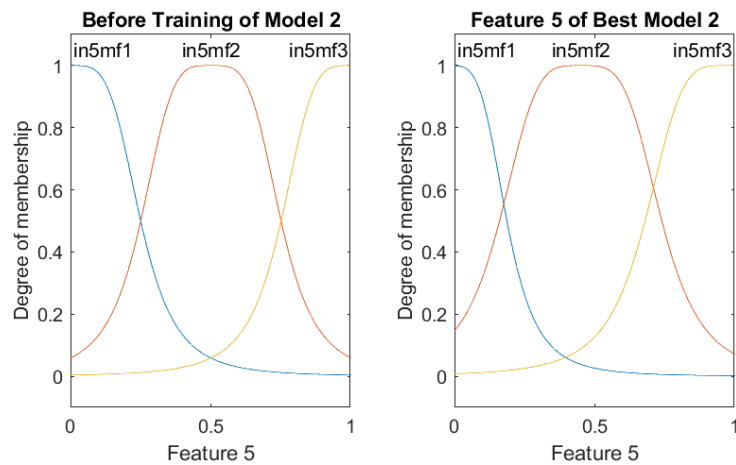
Διάγραμμα 9: TSK Model 2 – Feature 2



Διάγραμμα 10: TSK Model 2 – Feature 3

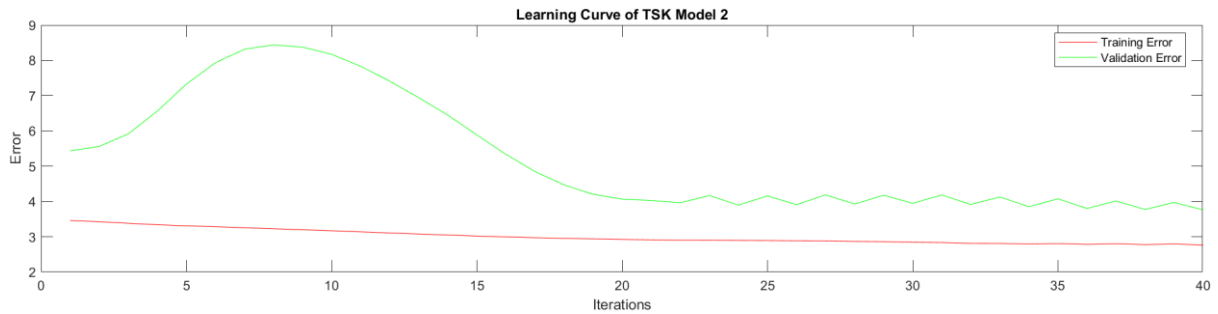


Διάγραμμα 11: TSK Model 2 – Feature 4

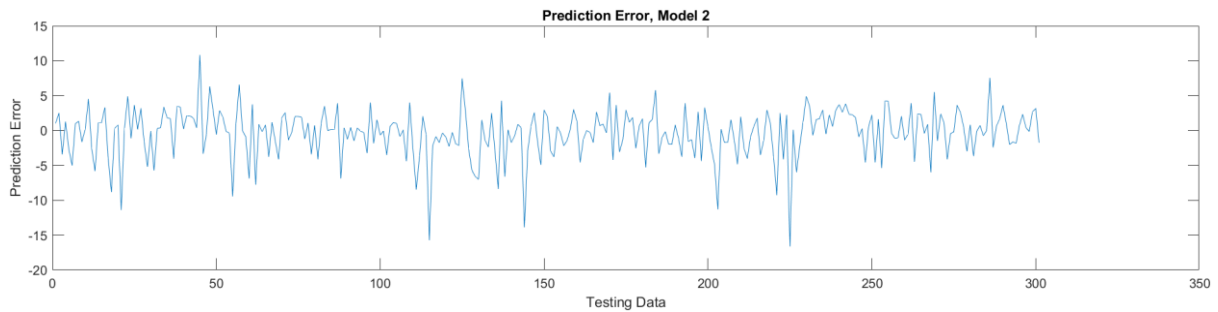


Διάγραμμα 12: TSK Model 2 – Feature 5

Καμπύλη Εκμάθησης και Σφάλμα Πρόβλεψης για το Μοντέλο 2:



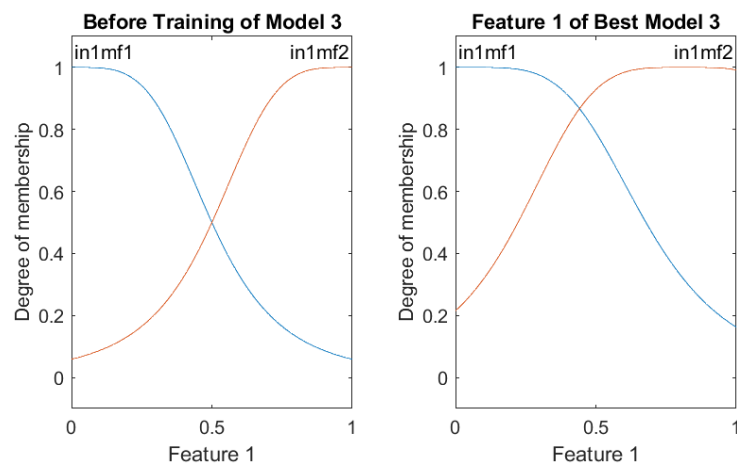
Διάγραμμα 13: TSK Model 2 – Καμπύλη Εκμάθησης



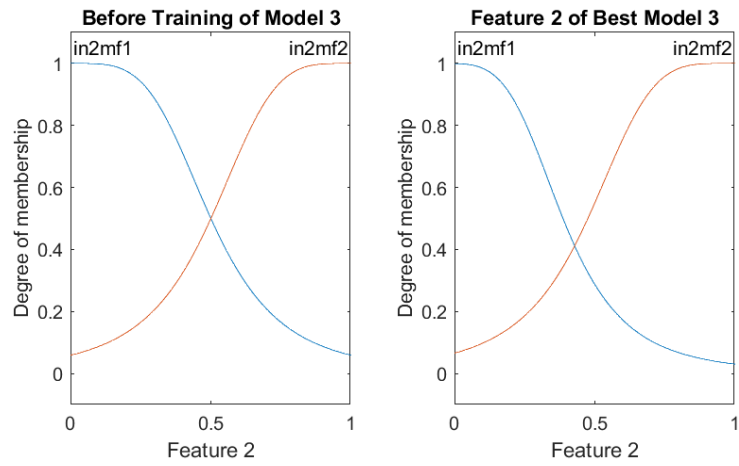
Διάγραμμα 14: TSK Model 2 – Σφάλμα Πρόβλεψης

Μοντέλο 3 (TSK_model_3)

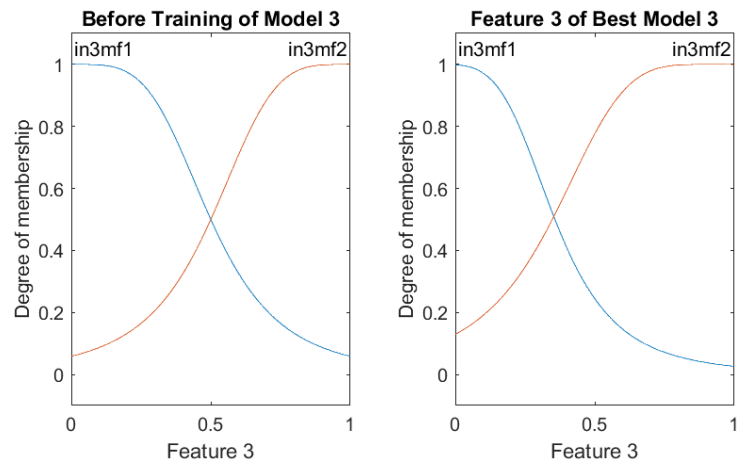
Διαγράμματα για κάθε Feature πριν και μετά την εκπαίδευση του Μοντέλου 3:



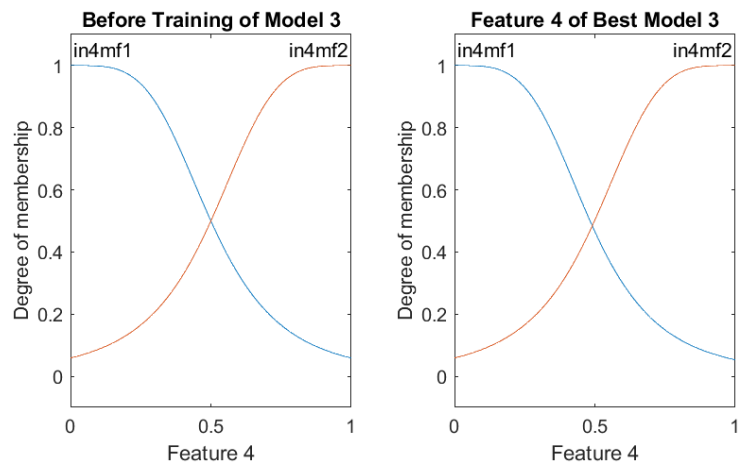
Διάγραμμα 15: TSK Model 3 – Feature 1



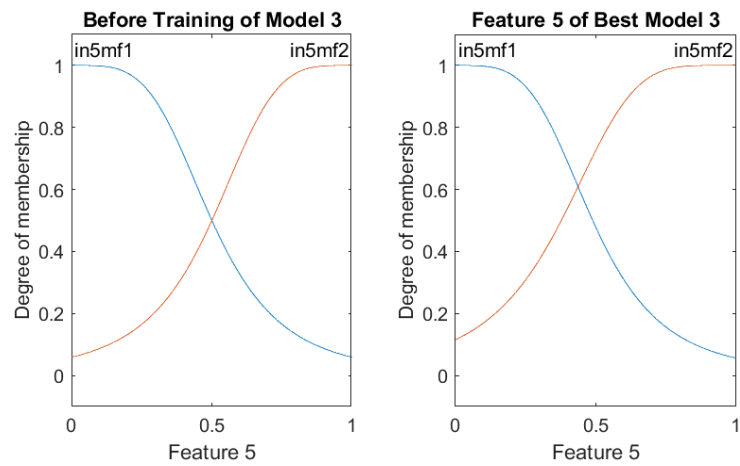
Διάγραμμα 16: TSK Model 3 – Feature 2



Διάγραμμα 17: TSK Model 3 – Feature 3

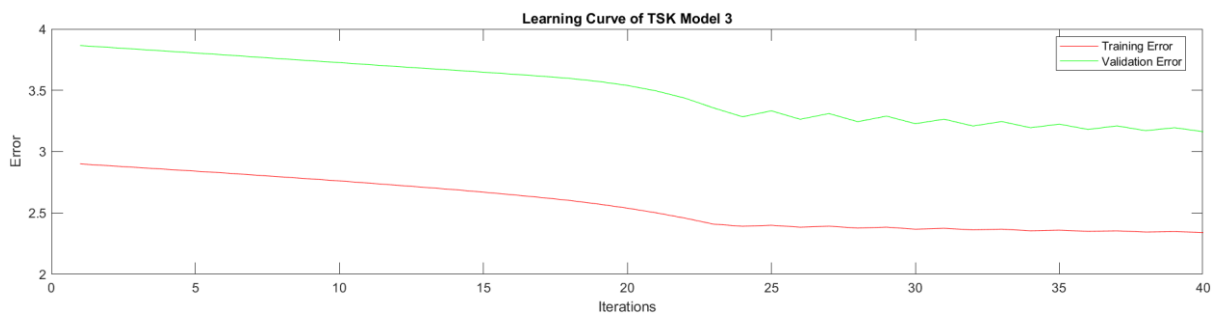


Διάγραμμα 18: TSK Model 3 – Feature 4



Διάγραμμα 19: TSK Model 3 – Feature 5

Καμπύλη Εκμάθησης και Σφάλμα Πρόβλεψης για το Μοντέλο 3:



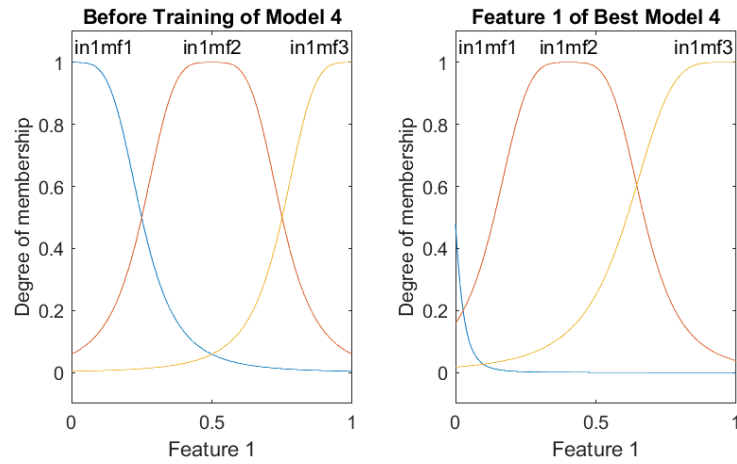
Διάγραμμα 20: TSK Model 3 – Καμπύλη Εκμάθησης



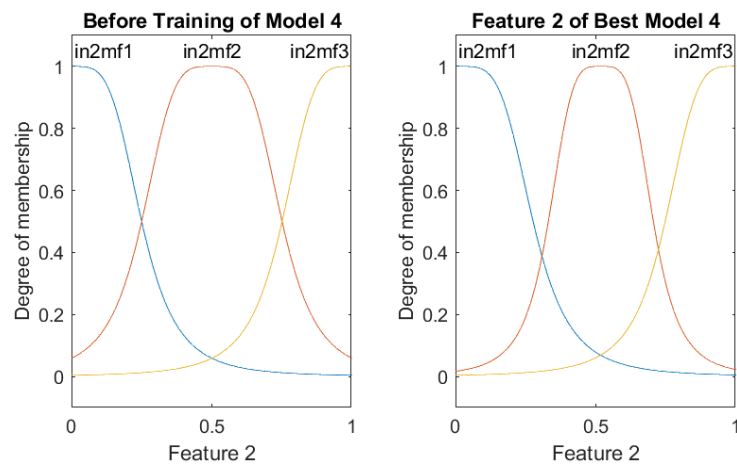
Διάγραμμα 21: TSK Model 3 – Σφάλμα Πρόβλεψης

Μοντέλο 4 (TSK_model_4)

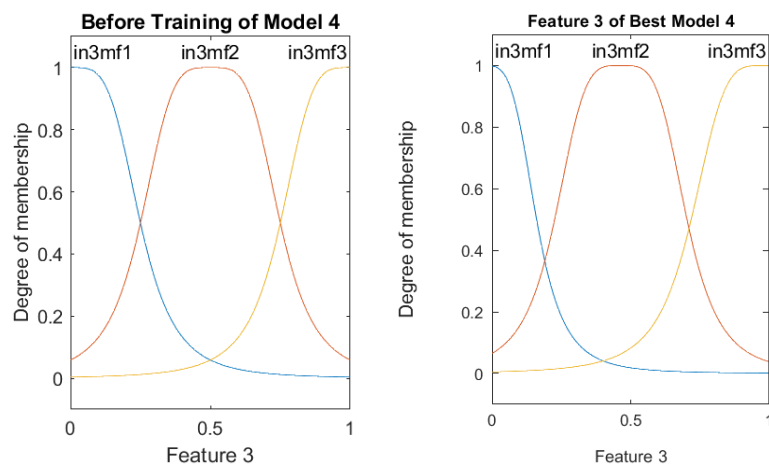
Διαγράμματα για κάθε Feature πριν και μετά την εκπαίδευση του Μοντέλου 4:



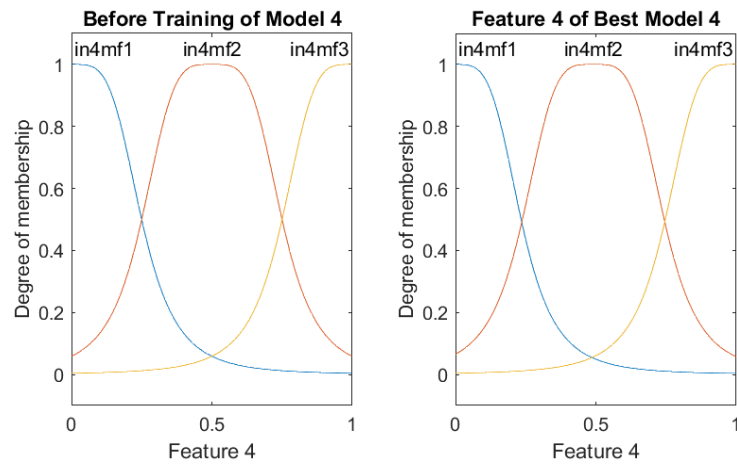
Διάγραμμα 22: TSK Model 4 – Feature 1



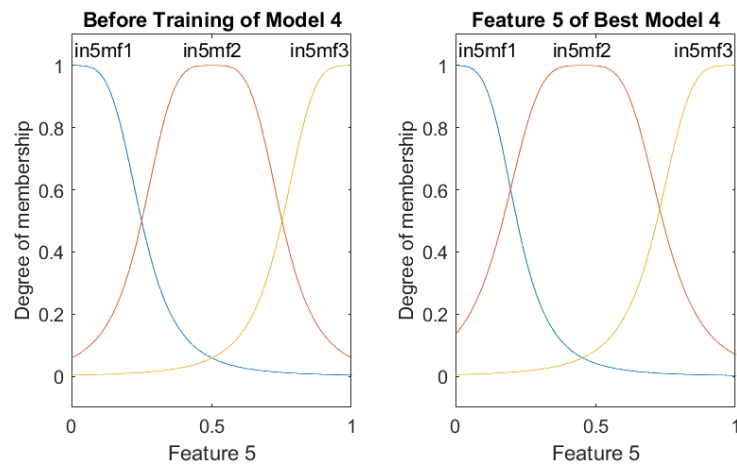
Διάγραμμα 23: TSK Model 4 – Feature 2



Διάγραμμα 24: TSK Model 4 – Feature 3

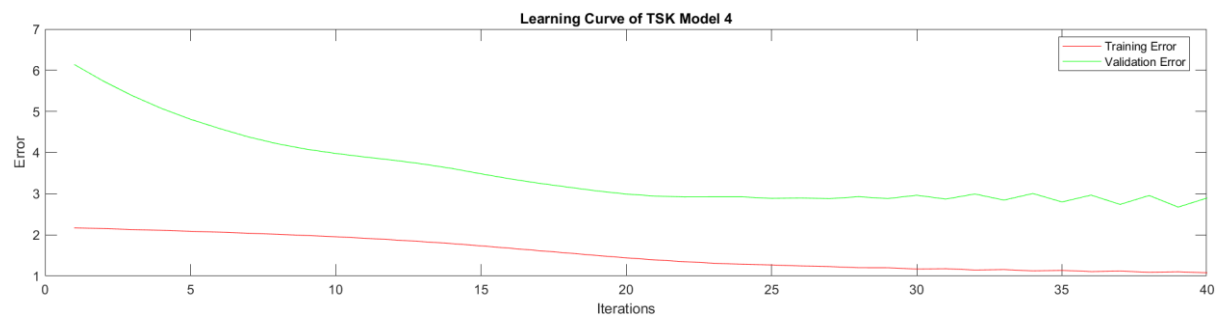


Διάγραμμα 25: TSK Model 4 – Feature 4



Διάγραμμα 26: TSK Model 4 – Feature 5

Καμπύλη Εκμάθησης και Σφάλμα Πρόβλεψης για το Μοντέλο 4:



Διάγραμμα 27: TSK Model 4 – Καμπύλη Εκμάθησης



Διάγραμμα 28: TSK Model 4 – Σφάλμα Πρόβλεψης

Τιμές Δεικτών Απόδοσης

Οι τιμές για τους δείκτες, με βάση τους οποίους αξιολογούνται τα μοντέλα TSK, δίνονται στον Πίνακα 2.

Μοντέλο	RMSE	NMSE	NDEI	R ²
TSK_Model_1	3.6228	0.2891	0.5377	0.7109
TSK_Model_2	3.5442	0.2767	0.5261	0.7233
TSK_Model_3	3.3399	0.2457	0.4957	0.7543
TSK_Model_4	2.7053	0.1612	0.4015	0.8388

Πίνακας 2: Τιμές Δεικτών Απόδοσης

Συμπεράσματα

Παρατηρώντας τόσο τα παραπάνω διαγράμματα, όσο και τις τιμές που αναγράφονται στον Πίνακα 2, μπορούν να γίνουν ορισμένες παρατηρήσεις και να εξαχθούν τα εξής συμπεράσματα:

- Τα polynomial μοντέλα (3 και 4) είναι κάνουν σαφώς καλύτερες προβλέψεις από τα singleton μοντέλα (1 και 2).
- Το Μοντέλο 4 έχει το μικρότερο σφάλμα, οπότε αυτό είναι και το καλύτερο μεταξύ των τεσσάρων. Γενικότερα, όσο μικρότερο το σφάλμα, τόσο καλύτερο το μοντέλο.
- Διακρίνεται φαινόμενο overfitting στα Μοντέλα 2 και 4, ενώ υπάρχει μια τάση προς αυτό και στο Μοντέλο 1, χωρίς όμως αυτό να είναι απολύτως ξεκάθαρο.

Δεύτερο Μέρος - Εφαρμογή σε Dataset με Υψηλή Διαστασιμότητα

Το δεύτερο σκέλος της εργασίας αφορά τη μελέτη ενός dataset με υψηλή διαστασιμότητα και, συνεπώς, αρκετά μεγαλύτερο όγκο δεδομένων από αυτό που εξετάστηκε κατά το πρώτο μέρος. Το γεγονός ότι ο αριθμός κανόνων αυξάνεται εκθετικά σε σχέση με το πλήθος των εισόδων για την περίπτωση του grid partitioning, οδηγεί στην αναζήτηση διαφορετικών μεθόδων για την μοντελοποίηση αυτού του προβλήματος. Προτεραιότητα, πλέον, αποτελεί η **μείωση της διαστασιμότητας**, κάτι που επιτυγχάνεται μέσω της **επιλογής χαρακτηριστικών**, καθώς και της **χρήσης διαμέρισης διασκορπισμού**.

Το σύνολο δεδομένων που θα μελετηθεί είναι το Superconductivity dataset από το UCI Repository, το οποίο περιλαμβάνει 21263 δείγματα (instances) και 81 χαρακτηριστικά (features). Ο κώδικας MATLAB που αποτελεί την υλοποίηση του μέρους αυτού της εργασίας είναι καταγεγραμμένος στο αρχείο “regression_b_10268.m”.

Αρχικά, όπως και στο πρώτο μέρος της εργασίας, γίνεται ο κατάλληλος διαχωρισμός του dataset, με σκοπό τον σχηματισμό τριών υποσυνόλων δεδομένων: του υποσυνόλου **εκπαίδευσης**, του υποσυνόλου **επικύρωσης** και του υποσυνόλου **ελέγχου απόδοσης** του τελικού μοντέλου (D_{chk}). Ο διαχωρισμός έγινε μέσω της συνάρτησης `split_scale` (αρχείο “split_scale.m”), ως εξής:

- Υποσύνολο Εκπαίδευσης (D_{trn}) → 60% του αρχικού dataset → 12758 instances
- Υποσύνολο Επικύρωσης (D_{val}) → 20% του αρχικού dataset → 4252 instances
- Υποσύνολο Ελέγχου (D_{chk}) → 20% του αρχικού dataset → 4253 instances

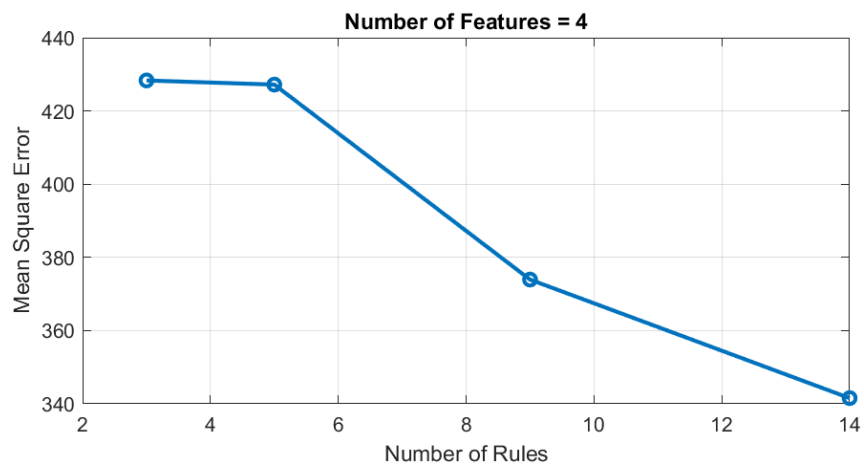
Όπως αναφέρεται και στην εκφώνηση, για τους σκοπούς της εργασίας έχουν οριστεί δύο βασικές παράμετροι: το **πλήθος των χαρακτηριστικών** που θα χρησιμοποιηθούν στην εκπαίδευση των μοντέλων και η **ακτίνα επιρροής** των clusters, η οποία επηρεάζει και το πλήθος των κανόνων. Για τις παραμέτρους αυτές, επιλέγονται οι τιμές που φαίνονται στον Πίνακα 3.

Τιμές Παραμέτρων				
Πλήθος Χαρακτηριστικών	4	8	12	16
Ακτίνα Επιρροής Clusters	0.3	0.5	0.8	1

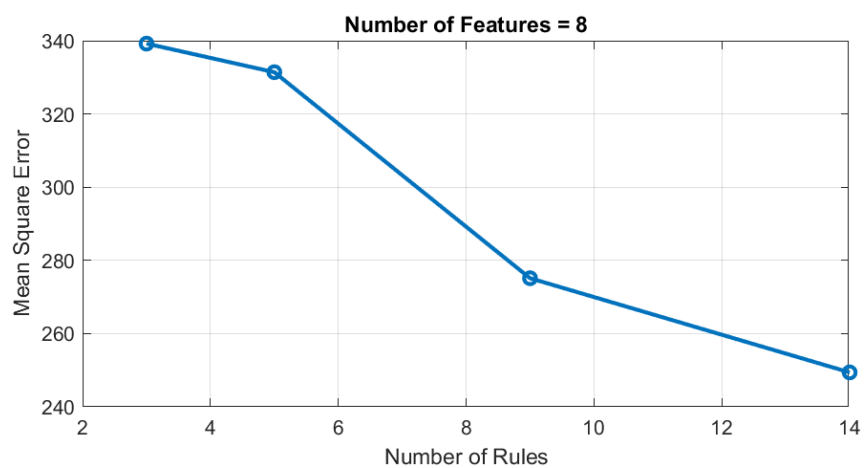
Πίνακας 3: Επιλογή Τιμών Παραμέτρων

Συνεπώς, θα εφαρμοστεί **5-fold-cross-validation** για κάθε μοντέλο που προκύπτει από τους συνδυασμούς των τιμών του Πίνακα 3, και τα μοντέλα αυτά θα αξιολογηθούν με βάση το μέσο τετραγωνικό σφάλμα.

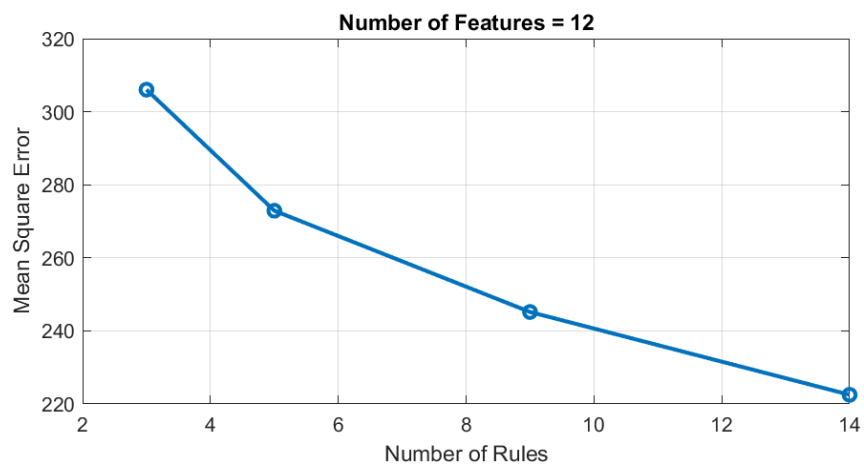
Διαγράμματα Σφάλματος Μοντέλων Συναρτήσεως του Πλήθους των Κανόνων που Χρησιμοποιήθηκαν:



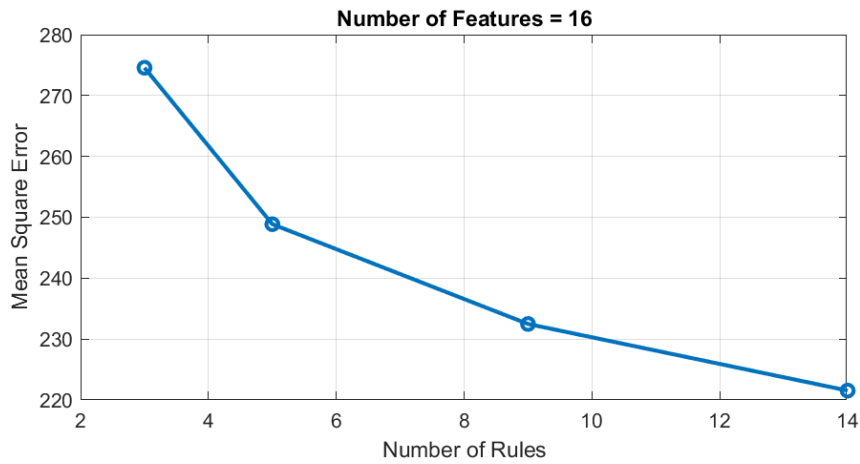
Διάγραμμα 29: MSE Συναρτήσεως του Πλήθους Κανόνων - 4 features



Διάγραμμα 30: MSE Συναρτήσεως του Πλήθους Κανόνων - 8 features

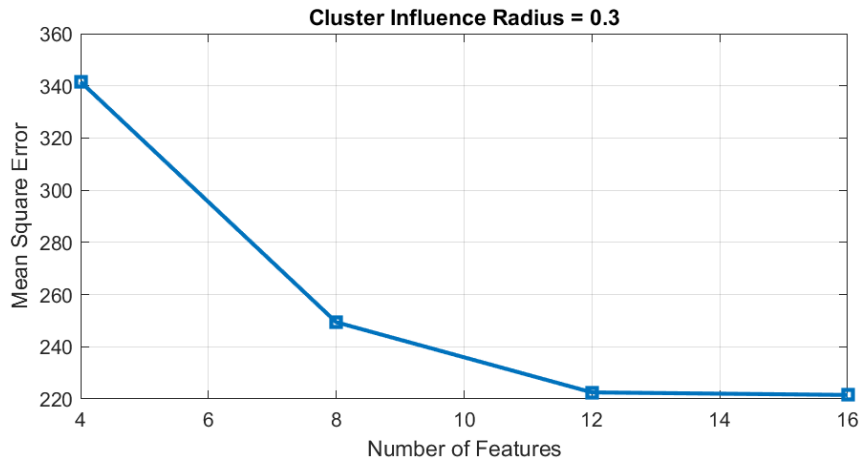


Διάγραμμα 31: MSE Συναρτήσεως του Πλήθους Κανόνων - 12 features

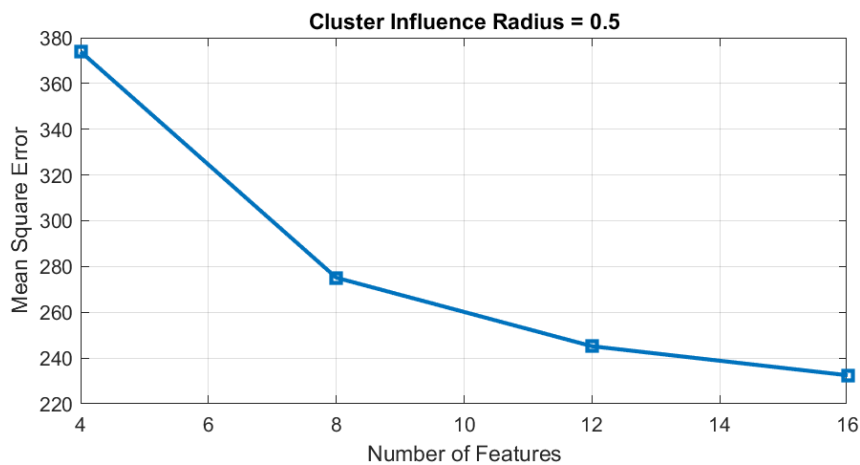


Διάγραμμα 32: MSE Συναρτήσεως του Πλήθους Κανόνων - 16 features

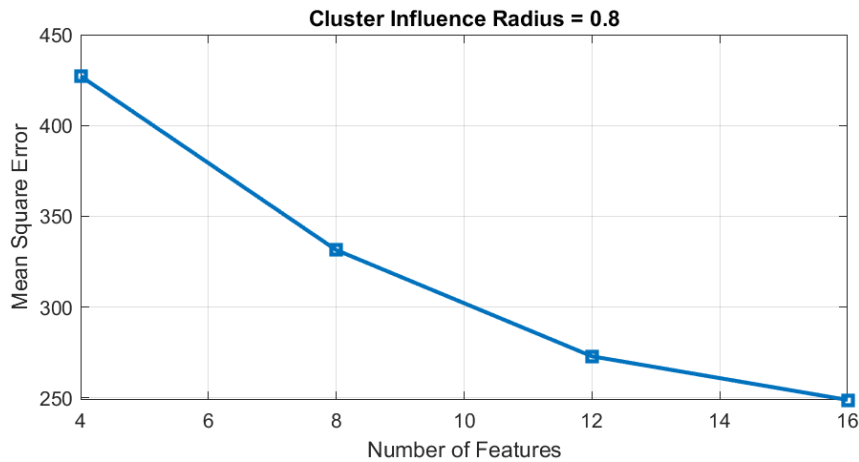
Διαγράμματα Σφάλματος Μοντέλων Συναρτήσεως του Πλήθους των Χαρακτηριστικών:



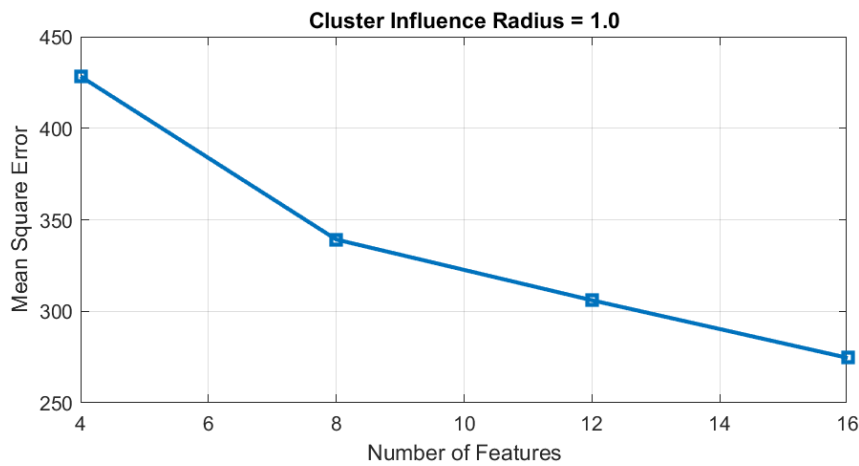
Διάγραμμα 33: MSE Συναρτήσεως του Πλήθους Χαρακτηριστικών - Ακτίνα = 0.3



Διάγραμμα 34: MSE Συναρτήσεως του Πλήθους Χαρακτηριστικών - Ακτίνα = 0.5



Διάγραμμα 35: MSE Συναρτήσεως του Πλήθους Χαρακτηριστικών - Ακτίνα = 0.8



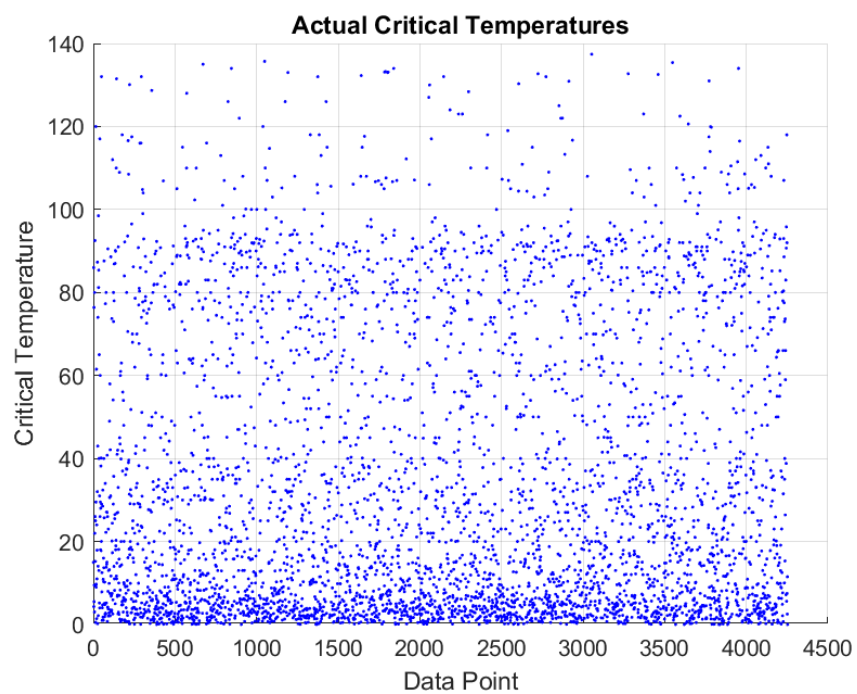
Διάγραμμα 36: MSE Συναρτήσεως του Πλήθους Χαρακτηριστικών - Ακτίνα = 1.0

Από τα παραπάνω διαγράμματα γίνεται εμφανές το γεγονός ότι όσο αυξάνεται το πλήθος των κανόνων, τόσο μειώνεται το σφάλμα της πρόβλεψης, κάτι που ισχύει και για τον αριθμό των χαρακτηριστικών. Επομένως, πρέπει να επιλεγεί ο κατάλληλος συνδυασμός πλήθους χαρακτηριστικών και ακτίνας, ώστε να επιτευχθεί μια αρκετά εύστοχη πρόβλεψη, ενώ παράλληλα να μην απαιτείται πολύ μεγάλη υπολογιστική ισχύς για την υλοποίηση του προγράμματος.

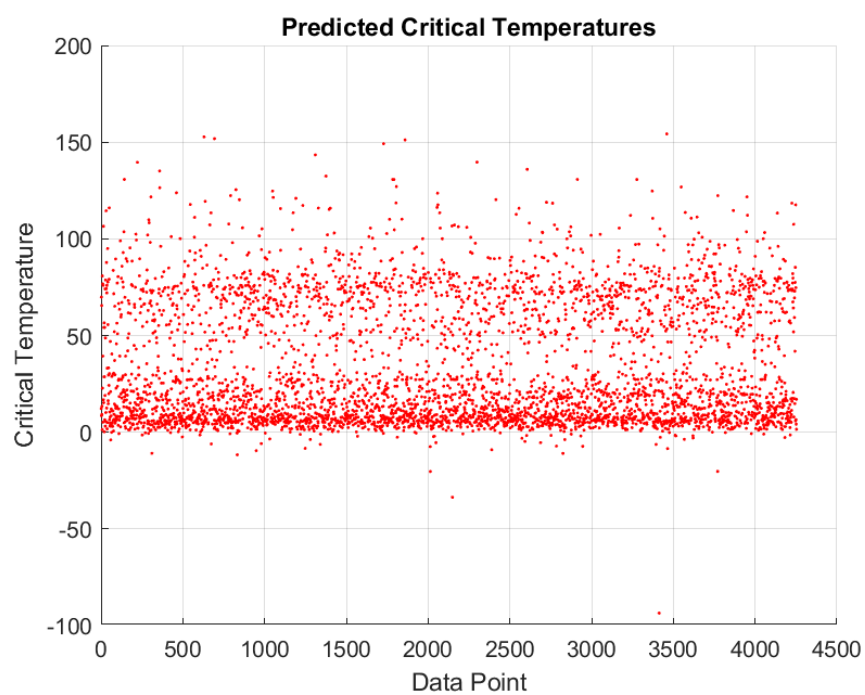
Εκτελώντας το πρόγραμμα που γράφτηκε για το συγκεκριμένο πρόβλημα, προκύπτει πως ο βέλτιστος συνδυασμός τιμών για τις δύο κύριες παραμέτρους, δηλαδή αυτός με τα μικρότερα σφάλματα, είναι:

- Πλήθος Χαρακτηριστικών = **16**
- Ακτίνα Επιρροής Clusters = **0.3**

Για το βέλτιστο μοντέλο, παρατίθενται τα εξής διαγράμματα:

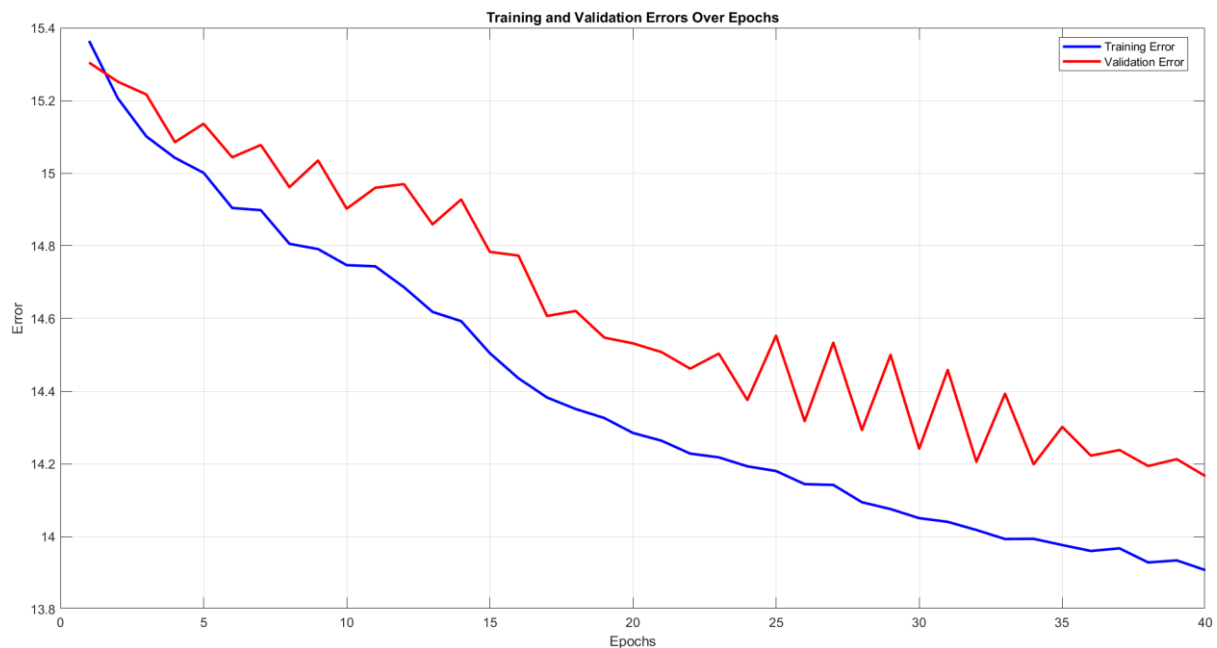


Διάγραμμα 37: Πραγματικές Τιμές Τελικού Μοντέλου

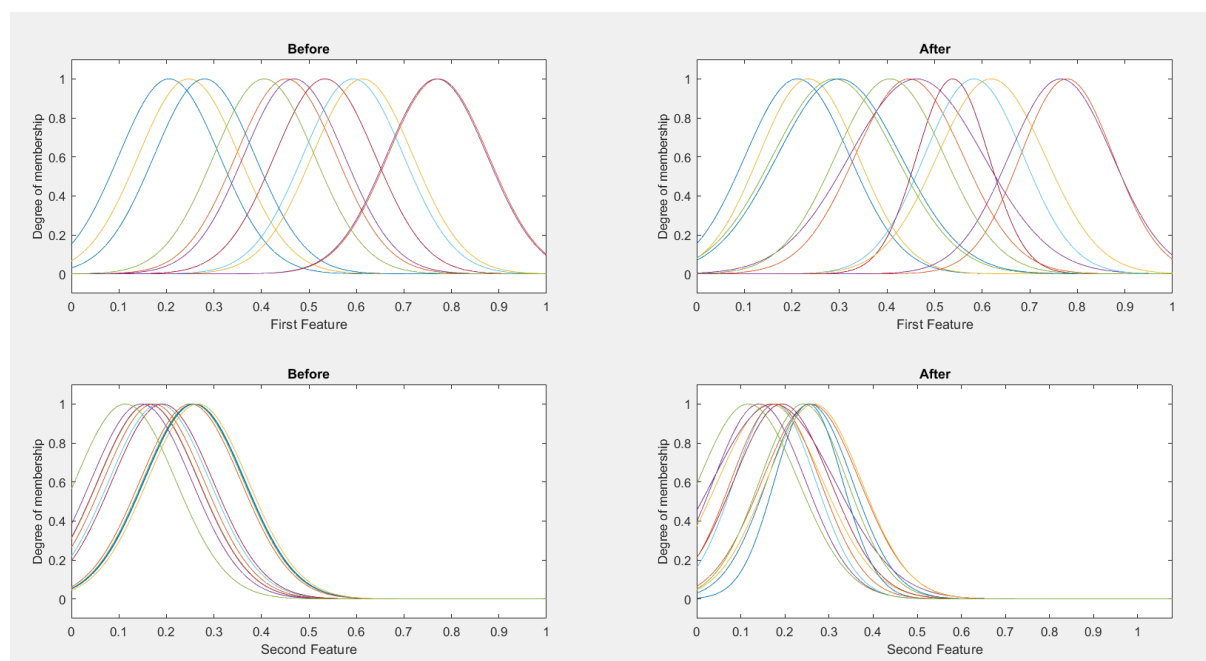


Διάγραμμα 38: Προβλέψεις Τελικού Μοντέλου

Με βάση τα δύο παραπάνω διαγράμματα, φαίνεται ότι η μεγαλύτερη πυκνότητα των προβλέψεων του τελικού μοντέλου είναι κοντά στο 0 και αραιώνουν προς τα πάνω, όπως συμβαίνει και στις πραγματικές τιμές. Ωστόσο, υπάρχουν και κάποιες προβλέψεις οι οποίες είναι εντελώς λάθος και έχουν τεράστια απόκλιση από τις πραγματικές τιμές.



Διάγραμμα 39: Σφάλματα Βέλτιστου Μοντέλου Συναρτήσεως Αριθμού Επαναλήψεων



Διάγραμμα 40: Συναρτήσεις Συμμετοχής Πριν και Μετά την Εκπαίδευση του Μοντέλου

Οι τιμές για τους δείκτες αξιολόγησης που ζητούνται, δίνονται στον Πίνακα 4.

Δείκτης Αξιολόγησης	RMSE	NMSE	NDEI	R ²
Τιμή	14.5437	0.1826	0.4273	0.8174

Πίνακας 4: Τιμές Δεικτών Απόδοσης Τελικού Μοντέλου

Συμπεράσματα

Βάσει των αποτελεσμάτων που προέκυψαν, είναι εμφανές πως η μέθοδος του Subtractive Clustering (SC) είναι πολύ αποδοτικότερη από την μέθοδο του Grid Partitioning που χρησιμοποιήθηκε στο πρώτο μέρος της εργασίας, καθώς οι προβλέψεις που παρέχει είναι αρκετά ικανοποιητικές, ενώ παράλληλα απαιτεί σημαντικά μικρότερο πλήθος κανόνων, γεγονός που, πρακτικά, εκμηδενίζει την απαιτούμενη υπολογιστική ισχύ για την εκτέλεσή της. Συγκεκριμένα, για το βέλτιστο μοντέλο της μεθόδου SC απαιτούνται μόλις τρεις ασαφείς κανόνες, ενώ για το Grid Partitioning θα απαιτούνταν 2^{16} ή 3^{16} κανόνες, για δύο και τρία ασαφή σύνολα ανά είσοδο, αντίστοιχα.

Σημείωση:

Κατά την υλοποίηση των προγραμμάτων τόσο του πρώτου, όσο και του δεύτερου μέρους της εργασίας, χρησιμοποιήθηκε σχετικά μικρός αριθμός επαναλήψεων (epochs), δηλαδή 40, χάριν συντομίας και εξοικονόμησης χρόνου κατά την εκτέλεση των προγραμμάτων. Προφανώς, για μεγαλύτερο αριθμό επαναλήψεων (π.χ. 100), τα αποτελέσματα θα ήταν αναλυτικότερα και πιο ξεκάθαρα.

Βιβλιογραφικές Πηγές

1. Γιάννης Μπουτάλης, Γεώργιος Συρακούσης, “Υπολογιστική Νοημοσύνη & Εφαρμογές”, Ξάνθη 2019