



Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης
Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Υπολογιστική Νοημοσύνη

8^ο Εξάμηνο

Επίλυση προβλήματος παλινδρόμησης με χρήση μοντέλων TSK

Σφυράκης Εμμανουήλ
AEM:9507
sfyrakise@ece.auth.gr

10 Μαρτίου 2021

Περιεχόμενα

1	Εφαρμογή σε απλό dataset	2
1.1	TSK_model_1	2
1.2	TSK_model_2	3
1.3	TSK_model_3	4
1.4	TSK_model_4	6
1.5	Σχολιασμός αποτελεσμάτων	7
2	Εφαρμογή σε dataset με υψηλή διαστασιμότητα	7

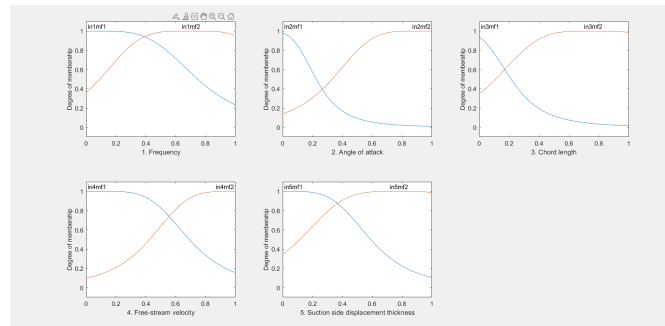
3	Εκπαίδευση βέλτιστου μοντέλου TSK	8
3.1	Σχολιασμός αποτελεσμάτων	9

1 Εφαρμογή σε απλό dataset

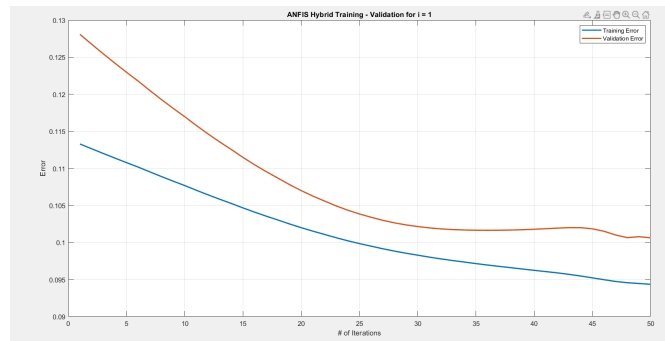
Στην παρούσα φάση, γίνεται η εκπαίδευση τεσσάρων απλών TSK μοντέλων επάνω στο Airfoil Self-Noise dataset του UCI repository. Τα ζητούμενα διαγράμματα και οι δείκτες απόδοσης του εκάστοτε μοντέλου παρουσιάζονται αναλυτικά στη συνέχεια.

1.1 TSK_model_1

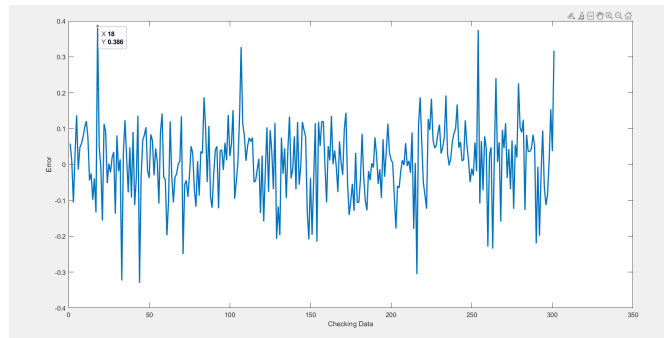
Το πρώτο μας μοντέλο αποτελείται από **2 συναρτήσεις συμμετοχής** ενώ στην έξοδο επιλέγεται ο **απο-ασαφοποιητής Singleton**. Τα διαγράμματα και οι δείκτες απόδοσης του μοντέλου δίνονται στα σχ.1.1, 1.2, 1.3 και στον πίνακα 1.



Σχήμα 1.1: Τελικές μορφές ασαφών συνόλων



Σχήμα 1.2: Διάγραμμα μάθησης σφάλματος-επαναλήψεων



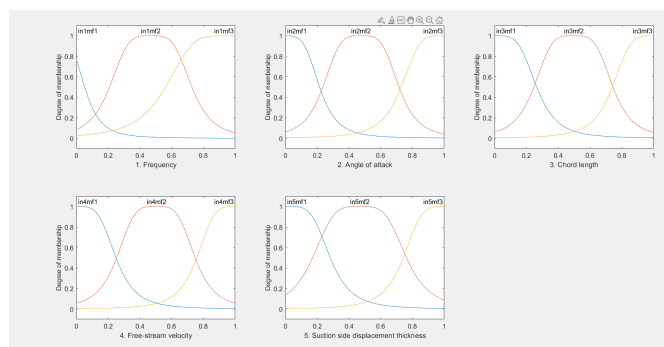
Σχήμα 1.3: Διάγραμμα σφάλματος πρόβλεψης

RMSE	NMSE	NDEI	R2
0.1036	0.3126	0.5591	0.6873

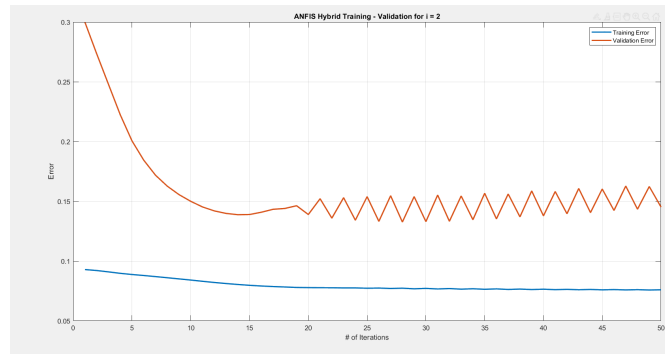
Πίνακας 1: Δείκτες απόδοσης

1.2 TSK_model_2

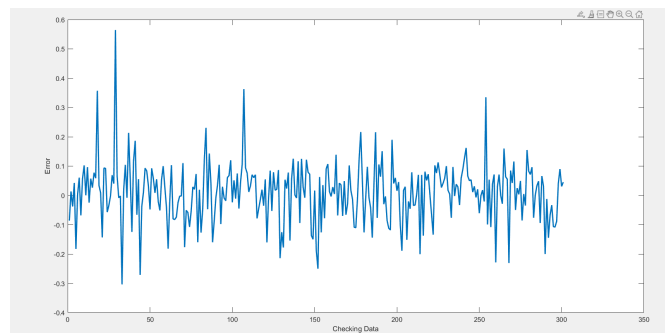
Το δεύτερο μοντέλο αποτελείται από **3 συναρτήσεις συμμετοχής** ενώ στην έξοδο επιλέγεται ο **απο-ασαφοποιητής Singleton**. Τα διαγράμματα και οι δείκτες απόδοσης του μοντέλου δίνονται στα σχ.1.4, 1.5, 1.6 και στον πίνακα 2.



Σχήμα 1.4: Τελικές μορφές ασαφών συνόλων



Σχήμα 1.5: Διάγραμμα μάθησης σφάλματος-επαναλήψεων



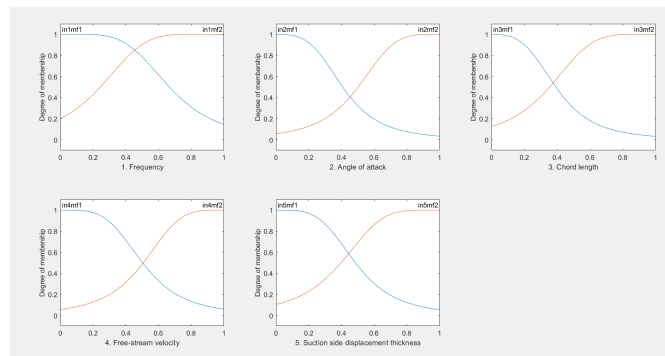
Σχήμα 1.6: Διάγραμμα σφάλματος πρόβλεψης

RMSE	NMSE	NDEI	R2
0.101	0.302	0.549	0.697

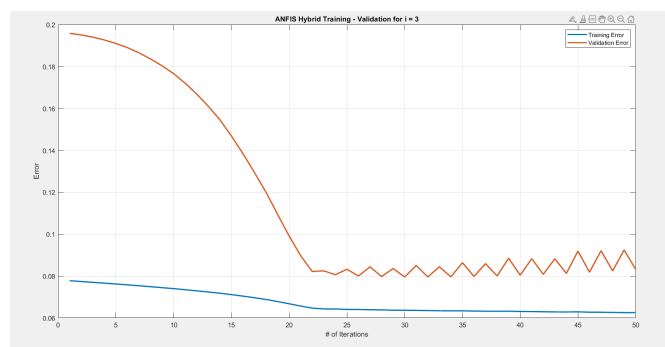
Πίνακας 2: Δείκτες απόδοσης

1.3 TSK_model_3

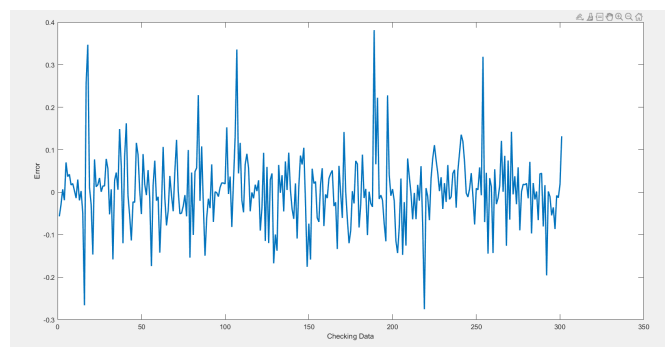
Το τρίτο μας μοντέλο αποτελείται από **2 συναρτήσεις συμμετοχής** ενώ στην έξοδο επιλέγεται ο **πολυωνυμικός απο-ασαφοποιητής**. Τα διαγράμματα και οι δείκτες απόδοσης του μοντέλου δίνονται στα σχ.1.7, 1.8, 1.9 και στον πίνακα 3.



Σχήμα 1.7: Τελικές μορφές ασαφών συνόλων



Σχήμα 1.8: Διάγραμμα μάθησης σφάλματος-επαναλήψεων



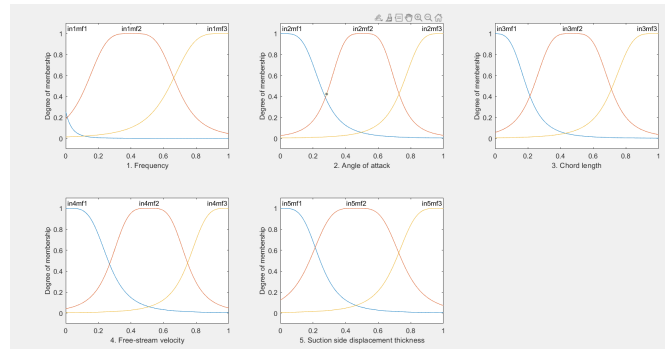
Σχήμα 1.9: Διάγραμμα σφάλματος πρόβλεψης

RMSE	NMSE	NDEI	R2
0.085	0.213	0.461	0.787

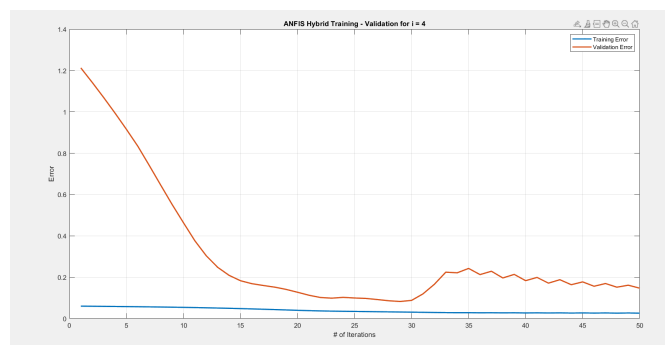
Πίνακας 3: Δείκτες απόδοσης

1.4 TSK_model_4

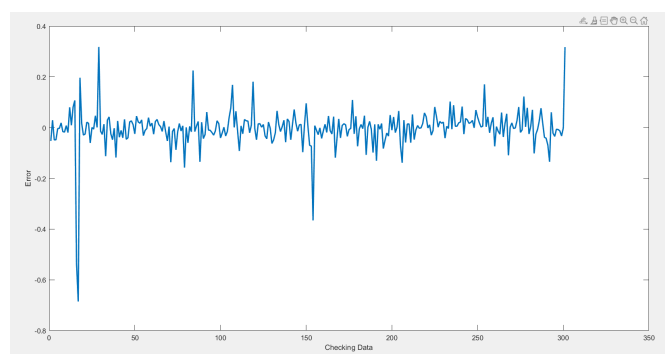
Το πρώτο μας μοντέλο αποτελείται από **3 συναρτήσεις συμμετοχής** ενώ στην έξοδο επιλέγεται ο **πολυωνυμικός απο-ασαφοποιητής**. Τα διαγράμματα και οι δείκτες απόδοσης του μοντέλου δίνονται στα σχ.1.10, 1.11, 1.12 και στον πίνακα 4.



Σχήμα 1.10: Τελικές μορφές ασαφών συνόλων



Σχήμα 1.11: Διάγραμμα μάθησης σφάλματος-επαναλήψεων



Σχήμα 1.12: Διάγραμμα σφάλματος πρόβλεψης

RMSE	NMSE	NDEI	R2
0.0787	0.18	0.424	0.819

Πίνακας 4: Δείκτες απόδοσης

1.5 Σχολιασμός αποτελεσμάτων

Γενικά, βλέπουμε ότι τα μοντέλα με **πολυωνυμική έξοδο** παρουσιάζουν **καλύτερη συμπεριφορά** όσον αφορά τους δείκτες απόδοσης. Βλέπουμε επίσης ότι τα μοντέλα με 3 συναρτήσεις συμμετοχής, είναι λίγο περισσότερο επιρρεπή σε φαινόμενα υπερεκπαίδευσης. Η παρατήρηση αυτή φαίνεται λογική καθώς όσο περισσότερες οι συναρτήσεις συμμετοχής, τόσο ευκολότερα τα μοντέλα εξειδικεύονται στο σετ εκπαίδευσης.

2 Εφαρμογή σε dataset με υψηλή διαστασιμότητα

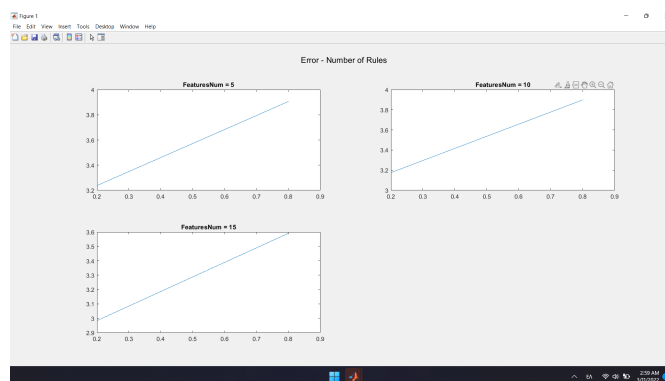
Στο δεύτερο κομμάτι της εργασίας, χρησιμοποιείται το Superconductivity dataset του UCI Repository, το οποίο θεωρείται dataset υψηλής διαστασιμότητας (81 χαρακτηριστικά). Οι ζητούμενες παράμετροι προς βελτιστοποίηση για την εκπαίδευση του μοντέλου είναι ο αριθμός των χαρακτηριστικών προς κράτηση(χρήση αλγορίθμου **relieff**) και η ακτίνα των clusters του αλγορίθμου **Subtractive Clustering** για την εύρεση του βέλτιστου πλήθους κανόνων IF-THEN.

Οι τιμές των παραμέτρων προς εξέταση είναι:

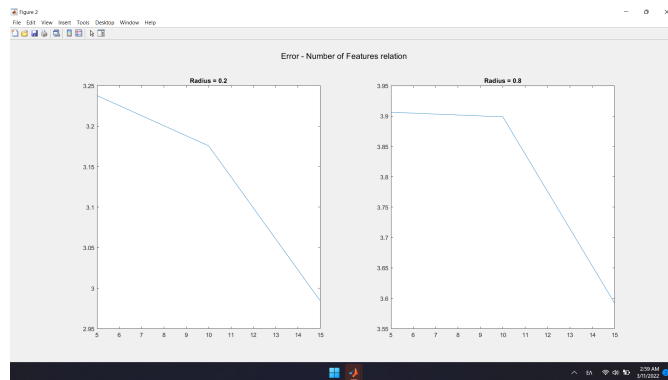
- Χαρακτηριστικά: [5, 10, 15]
- Ακτίνα clusters: [0.2, 0.8]

Να επισημάνουμε κάπου εδώ ότι δεν δόθηκαν περαιτέρω τιμές προς εξέταση στις παραμέτρους λόγω του μεγάλου χρόνου που χρειάζεται το Matlab για να τρέξει τον αλγόριθμο.

Η εκπαίδευση πραγματοποιήθηκε για 25 εποχές με τη βοήθεια της μεθόδου 5-fold cross validation. Από τα διαγράμματα 2.1, 2.2 παρατηρούμε ότι το μικρότερο δυνατό σφάλμα προκύπτει για $r_a = 0.2$ και αριθμό χαρακτηριστικών=15. Βλέπουμε ότι με την άυξηση της ακτίνας αυξάνεται το ελάχιστο μέσο σφάλμα ενώ το ίδιο συμβαίνει και με την μείωση των χαρακτηριστικών. Όσον αφορά βέβαια την ακτίνα, είναι λογικό κάτω από μια συγκεκριμένη τιμή να παρουσιαστεί υπερεκπαίδευση καθώς το μοντέλο θα εξειδικεύεται στα δεδομένα εκπαίδευσης.



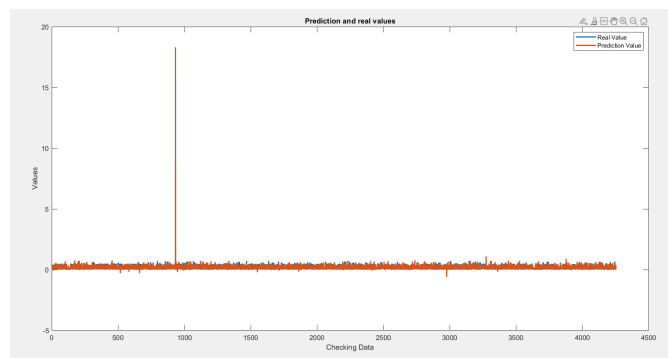
Σχήμα 2.1: Διάγραμμα σφάλματος τιμές ακτίνας clusters



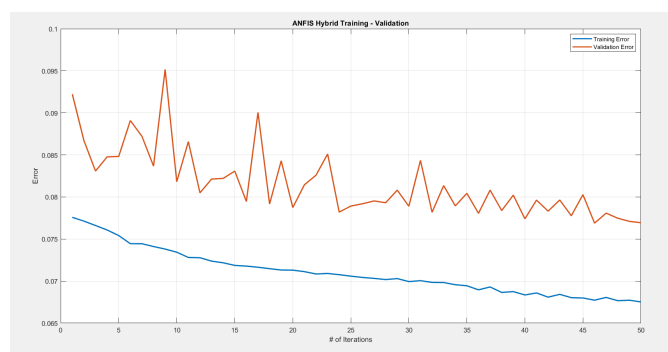
Σχήμα 2.2: Διάγραμμα σφάλματοςΑριθμού χαρακτηριστικών

3 Εκπαίδευση βέλτιστου μοντέλου TSK

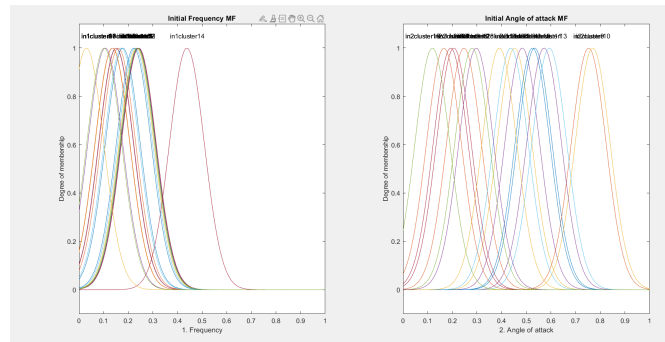
Εκπαιδύουμε το παραπάνω μοντέλο με τις βέλτιστες παραμέτρους για 50 εποχές. Τα ζητούμενα διαγράμματα και οι ζητούμενοι δείκτες απόδοσης δίνονται στα σχ. 3.1, 3.2, 3.3, 3.4 και στον πίνακα 5 αντίστοιχα.



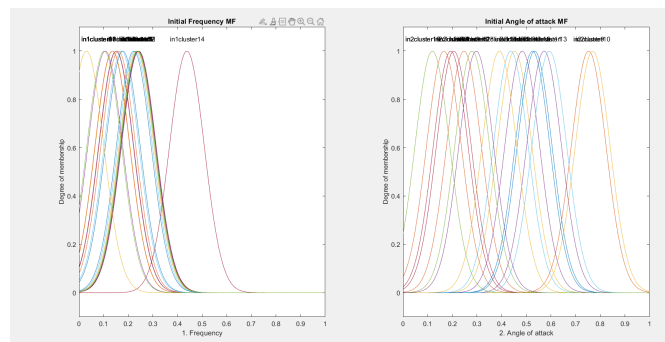
Σχήμα 3.1: Διάγραμμα πραγματικών τιμών και προβλέψεων μοντέλου



Σχήμα 3.2: Διάγραμμα σφάλματος-αριθμού επαναλήψεων



Σχήμα 3.3: Αρχική μορφή ασαφών συνόλων



Σχήμα 3.4: Τελική μορφή ασαφών συνόλων

RMSE	NMSE	NDEI	R2
0.2919	2.54	1.59	-1.547

Πίνακας 5: Δείκτες απόδοσης

3.1 Σχολιασμός αποτελεσμάτων

Από το διάγραμμα των προβλέψεων και των πραγματικών τιμών του μοντέλου βλέπουμε ότι το μοντέλο προβλέπει αρκετά καλά επάνω στα δεδομένα ελέγχου. Αν είχαμε επιλέξει την μέθοδο του Grid Partitioning με δύο ή τρία ασαφή σύνολα ανά είσοδο, θα είχαμε 2^{14} και 3^{14} κανόνες αντίστοιχα. Οι αριθμοί αυτοί είναι πάρα πολύ μεγάλοι συγκριτικά με τους 19 κανόνες του μοντέλου μας, άρα συμπεραίνουμε ότι θα ήταν υπερβολικά χρονοβόρο να χρησιμοποιήσουμε Grid Partitioning αντί για subtractive clustering.