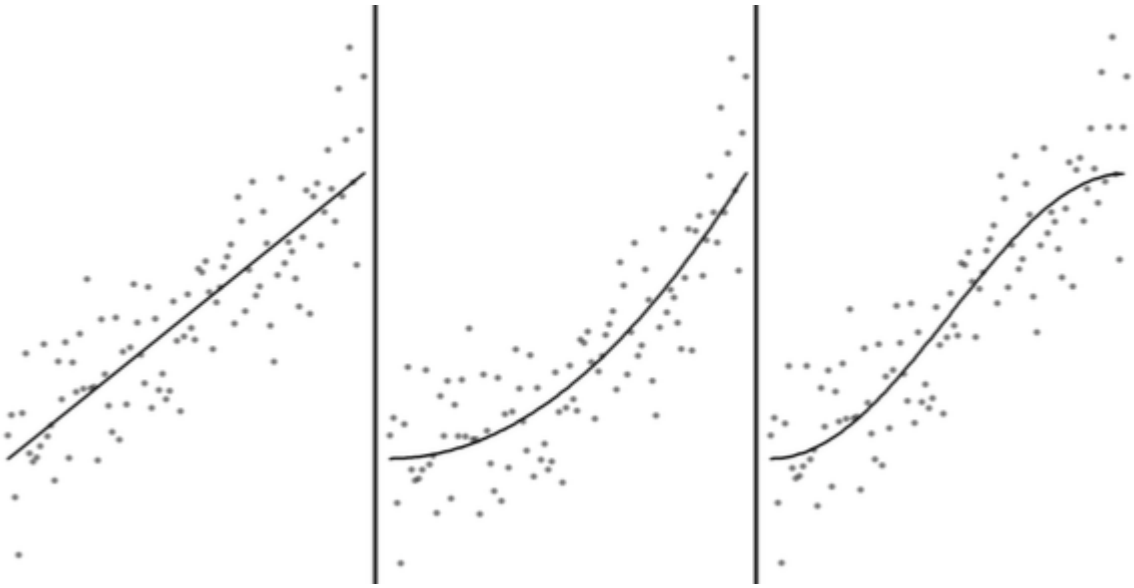
Report

REGRESSION ΜΕ ΧΡΗΣΗ ΜΟΝΤΕΛΩΝ TSK

ΟΝ/ΕΠ : ΝΙΚΟΛΑΟΣ ΙΣΤΑΤΙΑΔΗΣ

ΑΕΜ : 9175

Email : nikoista@ece.auth.gr



ΣΤΟΧΟΣ ΤΗΣ ΕΡΓΑΣΙΑΣ

Στόχος της εργασίας αυτής 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 βήματα προεπεξεργασίας όπως επιλογή χαρακτηριστικών (feature selection), καθώς και μεθόδους βελτιστοποίησης των μοντέλων μέσω της διασταυρωμένης επικύρωσης (cross validation).

1 ΕΦΑΡΜΟΓΗ ΣΕ ΑΠΛΟ DATASET

Στην πρώτη φάση της εργασίας, επιλέγεται από το UCI repository το Airfoil Self-Noise dataset, το οποίο περιλαμβάνει 1503 δείγματα (instances) και 6 χαρακτηριστικά (features). Ακολουθούμε τα εξής βήματα:

- Διαχωρισμός σε σύνολα εκπαίδευσης-επικύρωσης-ελέγχου:
  Σε πρώτη φάση είναι απαραίτητος ο διαχωρισμός του συνόλου δεδομένων σε τρία μη επικαλυπτόμενα υποσύνολα
  Dtrn-60%, Dval- 20%, Dchk-20% .

- Εκπαίδευση TSK μοντέλων με διαφορετικές παραμέτρους:
  Σε αυτό το στάδιο θα εξεταστούν διάφορα μοντέλα TSK όσον αφορά την απόδοσή τους στο σύνολο ελέγχου. Συγκεκριμένα, θα εκπαιδευτούν 4 TSK μοντέλα, στα οποία θα μεταβάλλονται η μορφή της εξόδου καθώς και το πλήθος των συναρτήσεων συμμετοχής για κάθε μεταβλητή εισόδου. Και τα 4 μοντέλα θα εκπαιδευτούν 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 πολυωνυμικής συνάρτησης εξόδου βελτιστοποιούνται μέσω της μεθόδου των ελαχίστων τετραγώνων (Least Squares). Οι συναρτήσεις συμμετοχής να είναι bell-shaped και η αρχικοποίησή τους να γίνει με τέτοιον τρόπο ώστε τα διαδοχικά ασαφή σύνολα να παρουσιάζουν σε κάθε είσοδο, βαθμό επικάλυψης περίπου 0.5.

	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.

- Αξιολόγηση μοντέλων: Για την ακρίβεια της εκτίμησης της πραγματικής συνάρτησης από κάθε ένα από τα παραπάνω μοντέλα, θα χρησιμοποιηθούν οι εξής δείκτες απόδοσης:

  1. MSE και RMSE

  2. Συντελεστής προσδιορισμού R^2 και $adjR^2$

  3. NMSE και NDEI

****

ΖΗΤΟΥΜΕΝΑ ΠΡΟΒΛΗΜΑΤΟΣ

Για κάθε ένα από τα 4 TSK μοντέλα που περιγράφονται στον παραπάνω πίνακα, να γίνουν οι κατάλληλες αρχικοποιήσεις και στη συνέχεια να εκτελεστεί η εκπαίδευση των μοντέλων με τις παραμέτρους που περιγράφτηκαν παραπάνω. Ως τελικό μοντέλο να επιλέγεται πάντα εκείνο το οποίο αντιστοιχεί στο μικρότερο σφάλμα στο σύνολο επικύρωσης . Για τις τέσσερις περιπτώσεις εκπαίδευσης .Να σχολιάσετε τα αποτελέσματα των μοντέλων τόσο όσον αφορά τη μορφή της εξόδου όσο και την διαμέριση του χώρου εισόδου. Το μεγαλύτερο πλήθος ασαφών συνόλων ανά είσοδο στην περίπτωση των αντίστοιχων TSK μοντέλων οδήγησε σε υπερεκπαίδευση· Να ερμηνευτούν οποιεσδήποτε διαφορές στην απόδοση των τεσσάρων μοντέλων.

Σύμφωνα με τα παραπάνω έχουμε τις εξής κατάλληλες αρχικοποιήσεις :

MATLAB CODE

inputMembershipFunction = " gbellmf ";

outputMembershipFunction = ["constant" , "constant" , "linear" , "linear"];

numberMembershipFunction = [2 , 3 , 2 , 3];

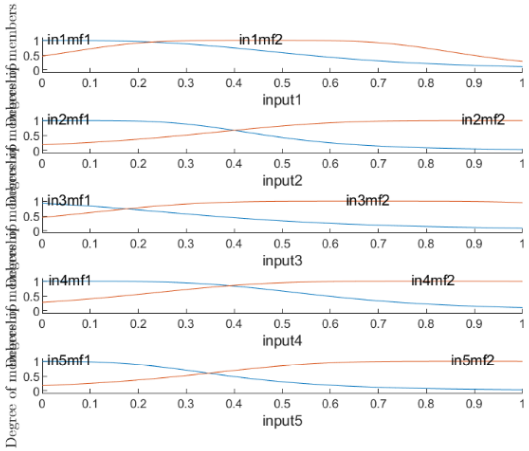
N = size(numberMembershipFunction,2);

Και στην συνέχεια εκπαιδεύεται το μοντέλο σύμφωνα με τα demo που έχουμε για τα TSK Μοντέλα στην σελίδα του μαθήματος αλλά και με μερικές ρυθμίσεις ακόμα.

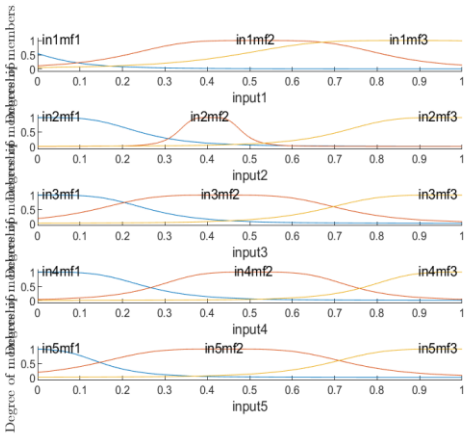
Για τα παρακάτω αποτελέσματα έχει εκτελεστεί το αρχείο
REGRESSION_1_SIMPLE .m

1) ΕΡΩΤΗΜΑ

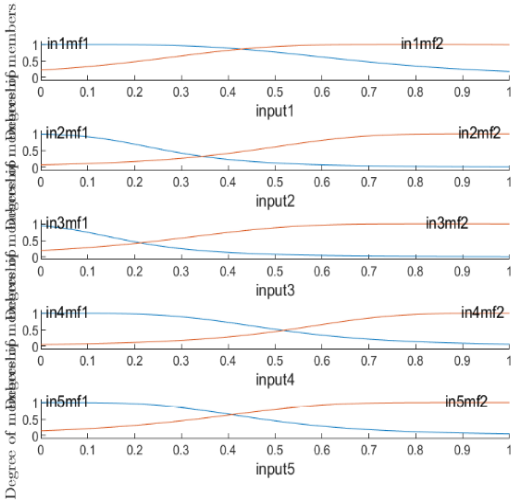
Να δώσετε τα αντίστοιχα διαγράμματα στα οποία να απεικονίζονται οι τελικές μορφές των ασαφών συνόλων που προέκυψαν μέσω της διαδικασίας εκπαίδευσης.



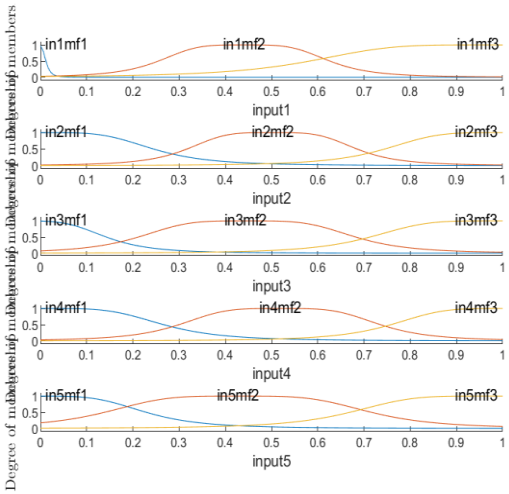
ΜΟΝΤΕΛΟ 1 ΕΙΣΟΔΟΙ



ΜΟΝΤΕΛΟ 2 ΕΙΣΟΔΟΙ



ΜΟΝΤΕΛΟ 3 ΕΙΣΟΔΟΙ

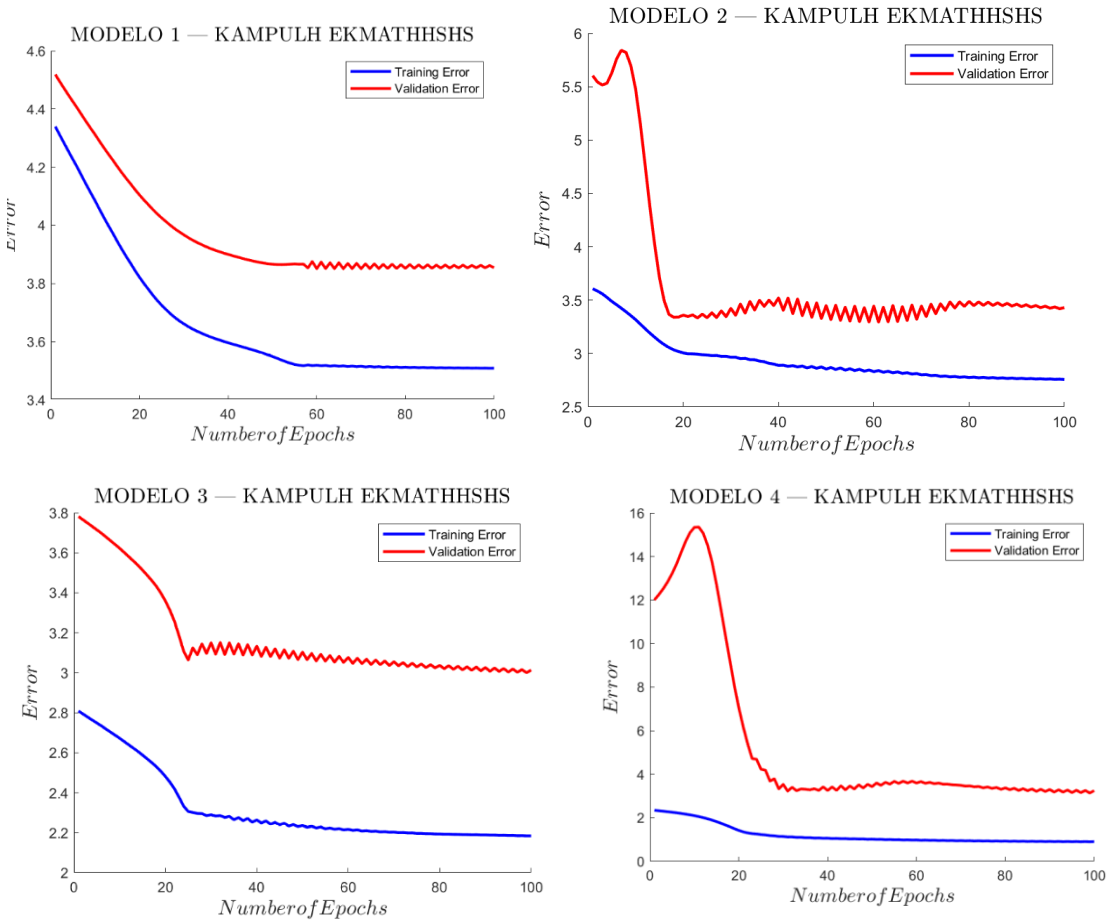


ΜΟΝΤΕΛΟ 4 ΕΙΣΟΔΟΙ

Ο συνολικός χώρος εισόδου διαμερίζετε πλήρως, ακολουθώντας στην φάση αυτή τον διαμερισμό πλέγματος (grid partition). Για κάθε είσοδο επιλέγεται ένας αριθμός ασαφών συνόλων, επιλέγοντας μια μορφή συνάρτησης συμμετοχής που στην συγκεκριμένη περίπτωση είναι **Singleton** και **Polynomial** έτσι ώστε δύο διαδοχικά ασαφή σύνολα να παρουσιάζουν μεταξύ τους έναν ικανοποιητικό βαθμό επικάλυψης (γύρω στο 0.5) . Είναι φανερό πως συμβαίνει αυτό σε μεγάλο βαθμό στις εισόδους (παρατηρούμε τα σημεία που τέμνονται)

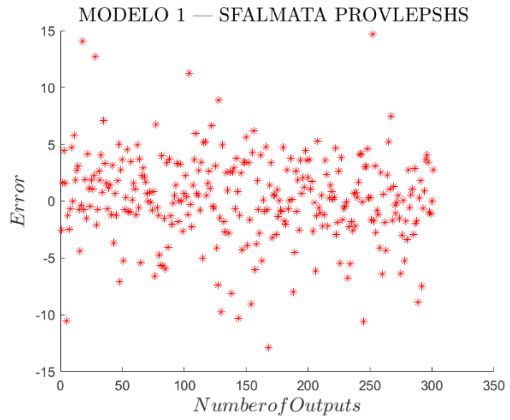
2) ΕΡΩΤΗΜΑ

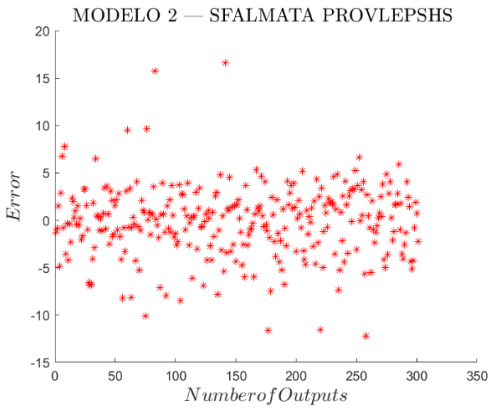
Να δοθούν τα διαγράμματα μάθησης (learning curves) όπου να απεικονίζεται το σφάλμα του μοντέλου συναρτήσει του αριθμού των επαναλήψεων (Iterations - Epochs).

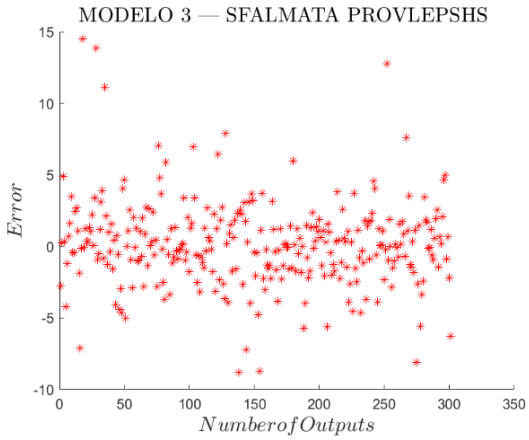


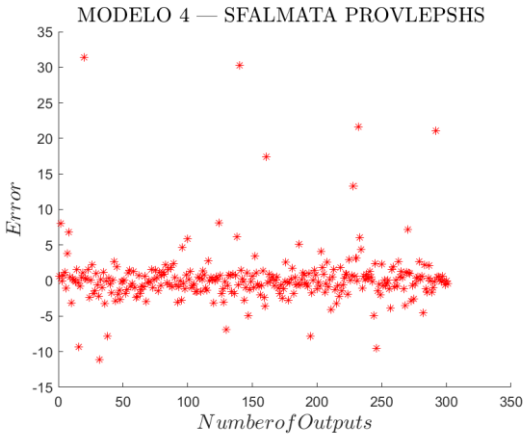
3) ΕΡΩΤΗΜΑ

Να δοθούν τα διαγράμματα όπου να αποτυπώνονται τα σφάλματα πρόβλεψης.









4) ΕΡΩΤΗΜΑ

Να παρουσιαστούν σε μορφή πίνακα οι τιμές των δεικτών απόδοσης RMSE, NMSE, NDEI, R^2

	Μοντέλο 1	Μοντέλο 2	Μοντέλο 3	Μοντέλο 4
R^2	0.7111	0.7027	0.8168	0.6340
RMSE	3.7510	3.7089	2.9870	3.9961
NMSE	0.2889	0.2973	0.1832	0.3660
NDEI	0.5375	0.5452	0.4280	0.6050

ΣΧΟΛΙΑΣΜΟΣ ΑΠΟΤΕΛΕΣΜΑΤΩΝ

Συγκρίνοντας το R^2 βλέπουμε πως το Μοντέλο 3 είναι το καλύτερο μετά το Μοντέλο 1 και μετά το Μοντέλο 2 και τέλος το 4.

Είναι φανερό πως το μεγαλύτερο πλήθος ασαφών συνόλων ανά είσοδο στην περίπτωση των αντίστοιχων TSK μοντέλων οδήγησε σε **υπερεκπαίδευση**. Αυτό ερμηνεύουν και τα learning curves αλλά και τα διαγράμματα σφαλμάτων προβλέψεις όπου στα Μ**οντέλα 2,4** . Επίσης βλέπουμε πως το Μοντέλο 3 έχει το πιο χαμηλούς δείκτες RMSE, NMSE, NDEI σε σχέση με τα άλλα 4 Μοντέλα. Από την άλλη πλευρά το Μοντέλο 4 όπου έχουμε πολυωνυμική είσοδο (3) έχει τους υψηλούς δείκτες RMSE, NMSE, NDEI και το πιο χαμηλό R^2 κανοντάς το το πιο ακατάλληλο μοντέλο.

Ακολουθώντας το θεωρητικό υπόβαθρο των διαλέξεων θα γίνει συμπερασματολογία πάνω στις γραφικές παραστάσεις της διαδικασίας εκπαίδευσης του δικτύου, οπότε έχουμε τα εξής:

ΜΟΝΤΕΛΟ 1

Στο μοντέλο αυτό έχουμε επιτυχή διαδικασία εκπαίδευσης του δικτύου διότι το MSEtrn μειώνεται συναρτήσει των διαδοχικών επαναλήψεων εκπαίδευσης και προσεγγίζει μια σταθερή τιμή περίπου στο 3.5 . Επίσης το MSEchk μειώνεται συναρτήσει των διαδοχικών επαναλήψεων εκπαίδευσης και προσεγγίζει μια σταθερή τιμή περίπου στο 3.9 . Άρα στο μοντέλο 1 το TrainingError και το ValidationError έχουν μια διαφορά περίπου 0.4 . Σε αυτό το μοντέλο δεν εμφανίζεται το φαινόμενο τις υπερεκπαίδευσης. Το μοντέλο έχει καλή απόδοση μιας και το $R^2 = 0.7111$ δείχνοντας έτσι την ικανότητα του για προβλέψεις.

ΜΟΝΤΕΛΟ 2

Στο μοντέλο αυτό έχουμε ένα ανεπιθύμητο φαινόμενο κατά την εκπαίδευση, αυτό της υπερ-εκπαίδευσης όχι σε τόσο μεγάλο βαθμό όσο σε σχέση με το μοντέλο 4 (over-training). Το φαινόμενο αυτό εμφανίζεται όταν το δίκτυο προσαρμόζεται σε μεγάλο βαθμό στα δεδομένα του συνόλου εκπαίδευσης Dtrn , επιτυγχάνοντας πολύ χαμηλές τιμές του MSEtrn προς στο τέλος. Σε αυτή την περίπτωση, από ένα σημείο και μετά (επανάληψη 57), το σφάλμα ελέγχου MSEchk αρχίζει να αποκλίνει, πράγμα που σηματοδοτεί ότι το δίκτυο χάνει την ικανότητα γενίκευσης. Αυτό σημαίνει ότι η διαδικασία εκπαίδευσης θα πρέπει να τερματιστεί στο σημείο (επανάληψη 57), όπου το δίκτυο επιτυγχάνει μια ικανοποιητική ισορροπία ανάμεσα στην ακρίβεια προσέγγισης και στην γενίκευση, δηλαδή την ιδιότητα επιτυχούς πρόβλεψης της εξόδου σε άγνωστες εισόδους στο μοντέλο. Το μοντέλο έχει καλή απόδοση μιας και το $R^2 = 0.7027$ δείχνοντας έτσι την ικανότητα του για προβλέψεις.

ΜΟΝΤΕΛΟ 3

Στο μοντέλο αυτό είναι το καλύτερο από τα 4 καθώς έχουμε επιτυχή διαδικασία εκπαίδευσης του δικτύου διότι το MSEtrn μειώνεται συναρτήσει των διαδοχικών επαναλήψεων εκπαίδευσης και προσεγγίζει μια σταθερή τιμή περίπου στο 2.2 . Επίσης το MSEchk μειώνεται συναρτήσει των διαδοχικών επαναλήψεων εκπαίδευσης και προσεγγίζει μια σταθερή τιμή περίπου στο 3. Άρα στο μοντέλο 3 το TrainingError και το ValidationError έχουν μια διαφορά περίπου 1.2 . Σε αυτό το μοντέλο δεν εμφανίζεται το φαινόμενο τις υπερεκπαίδευσης και η απόδοσή του είναι η καλύτερη σε σχέση μετά άλλα μοντέλα σύμφωνα με το $R^2 = 0.8168$

ΜΟΝΤΕΛΟ 4

Όπως μιλήσαμε και πιο πάνω Στο μοντέλο αυτό έχουμε ένα ανεπιθύμητο φαινόμενο κατά την εκπαίδευση, αυτό της υπερ-εκπαίδευσης (over-training). Το φαινόμενο αυτό εμφανίζεται όταν το δίκτυο προσαρμόζεται σε μεγάλο βαθμό στα δεδομένα του συνόλου εκπαίδευσης Dtrn , επιτυγχάνοντας πολύ χαμηλές τιμές του MSEtrn προς στο τέλος . Σε αυτή την περίπτωση, από ένα

σημείο και μετά (επανάληψη 25), το σφάλμα ελέγχου MSEchk αρχίζει να αποκλίνει, πράγμα που σηματοδοτεί ότι το δίκτυο χάνει την ικανότητα γενίκευσης. Αυτό σημαίνει ότι η διαδικασία εκπαίδευσης θα πρέπει να τερματιστεί στο σημείο (επανάληψη 25), όπου το δίκτυο επιτυγχάνει μια ικανοποιητική ισορροπία ανάμεσα στην ακρίβεια προσέγγισης και στην γενίκευση, δηλαδή την ιδιότητα επιτυχούς πρόβλεψης της εξόδου σε άγνωστες εισόδους στο μοντέλο.

****

2 ΕΦΑΡΜΟΓΗ ΣΕ DATASET ΜΕ ΥΨΗΛΗ ΔΙΑΣΤΑΣΙΜΟΤΗΤΑ

Στη δεύτερη φάση της εργασίας θα ακολουθηθεί μια πιο συστηματική προσέγγιση στο πρόβλημα μοντελοποίησης μιας άγνωστης συνάρτησης. Για το σκοπό αυτό θα επιλεχθεί ένα dataset με υψηλότερο βαθμό διαστασιμότητας. ΄Ενα προφανές πρόβλημα που ανακύπτει από την επιλογή αυτή, είναι η λεγόμενη “έκρηξη” του πλήθους των IF-THEN κανόνων (rule explosion).

Όπως είναι γνωστό από τη θεωρία, για την κλασική περίπτωση του grid-partitioning του χώρου εισόδου, ο αριθμός των κανόνων αυξάνεται εκθετικά σε σχέση με το πλήθος των εισόδων, γεγονός που καθιστά πολύ δύσκολη την μοντελοποίηση μέσω ενός TSK μοντέλου ακόμα και για datasets μεσαίας κλίμακας.

Το dataset που θα επιλεγεί για την επίδειξη των παραπάνω μεθόδων είναι το Superconductivty dataset από το UCI Repository, το οποίο περιλαμβάνει 21263 δείγματα καθένα από τα οποία περιγράφεται από 81 μεταβλητές/χαρακτηριστικά. Είναι φανερό ότι το μέγεθος του dataset καθιστά απαγορευτική μια απλή εφαρμογή ενός ΤΣΚ μοντέλου, σαν αυτή του προηγούμενου μέρους της εργασίας. Ο μεγάλος αριθμός μεταβλητών καθιστά αναγκαία τη χρήση μεθόδων μείωσης της διαστασιμότητας καθώς και του αριθμού των IF-THEN κανόνων (π.χ. με 81 μεταβλητές/predictors, διαμερίζαμε το χώρο εισόδου κάθε μεταβλητής με δύο ασαφή σύνολα, θα καταλήγαμε με 2 81 κανόνες).

Ο στόχος αυτός θα επιτευχθεί μέσω της επιλογής χαρακτηριστικών και της χρήσης dispersion partition. Οι δύο αυτές μέθοδοι όμως, παρά τη ελάττωση της πολυπλοκότητας που επιφέρουν, εισάγουν στο πρόβλημα δύο ελεύθερες παραμέτρους, συγκεκριμένα:

- τον αριθμό των χαρακτηριστικών προς επιλογή
- και τον αριθμό των ομάδων που θα δημιουργηθούν.

Η επιλογή των δύο αυτών παραμέτρων επαφίεται στον εκάστοτε χρήστη και είναι ουσιαστική όσον αφορά την τελική απόδοση του μοντέλου. Στην παρούσα εργασία, θα υλοποιηθεί η μέθοδος αναζήτησης πλέγματος για την εύρεση των βέλτιστων τιμών των παραμέτρων.

Αναλυτικά, η μοντελοποίηση του προβλήματος θα ακολουθήσει λοιπόν τα εξής βήματα:

- Διαχωρισμός σε σύνολα εκπαίδευσης-επικύρωσης-ελέγχου:
  Σε πρώτη φάση είναι απαραίτητος ο διαχωρισμός του συνόλου δεδομένων σε τρία μη επικαλυπτόμενα υποσύνολα
  Dtrn-60%, Dval- 20%, Dchk-20% .
- Επιλογή των βέλτιστων παραμέτρων: Όπως αναφέρθηκε παραπάνω, το σύστημά μας περιλαμβάνει **δύο ελεύθερες παραμέτρους** την τιμή των οποίων πρέπει να επιλέξουμε εμείς. Η δημοφιλέστερη μέθοδος μέσω της οποίας επιτυγχάνεται αυτό είναι η **αναζήτηση πλέγματος**. Συγκεκριμένα, αφού λάβουμε ένα σύνολο τιμών για κάθε παράμετρο, δημιουργούμε ένα **n-διάστατο πλέγμα** (στην περίπτωσή μας **n = 2**), όπου κάθε σημείο αντιστοιχεί σε μια n-άδα τιμών για τις εν λόγω παραμέτρους, και σε κάθε σημείο χρησιμοποιούμε μια **μέθοδο αξιολόγησης** για ελέγξουμε την ορθότητα των συγκεκριμένων τιμών. Μια καθιερωμένη επιλογή για την αξιολόγηση αυτή αποτελεί η διασταυρωμένη επικύρωση **(cross validation)**.

- Σύμφωνα με τη μέθοδο αυτή ,και για επιλεγμένες τιμές των παραμέτρων, χωρίζουμε το σύνολο εκπαίδευσης σε δύο υποσύνολα, από τα οποία το ένα θα χρησιμοποιηθεί για την εκπαίδευση ενός μοντέλου και το δεύτερο για την αξιολόγησή του. Η διαδικασία αυτή επαναλαμβάνεται – συνήθως πέντε ή δέκα φορές – όπου κάθε φορά χρησιμοποιείται διαφορετικός διαχωρισμός του συνόλου εκπαίδευσης, και στο τέλος λαμβάνουμε τον μέσο όρο του σφάλματος του μοντέλου. Η λογική πίσω από τις πολλαπλές εκπαιδεύσεις και ελέγχους έγκειται στο ότι με αυτό τον τρόπο, αποκτούμε μια αρκετά καλή εκτίμηση της απόδοσης του μοντέλου, και έμμεσα των τιμών των παραμέτρων με βάση τις οποίες χτίστηκε το μοντέλο. Όταν η παραπάνω διαδικασία εκτελεστεί για κάθε σημείο του πλέγματος, λαμβάνουμε ως βέλτιστες τιμές των παραμέτρων, τις τιμές που αντιστοιχούν στο μοντέλο που παρουσίασε το ελάχιστο μέσο σφάλμα. Οι τιμές αυτές χρησιμοποιούνται για την εκπαίδευση του τελικού μας μοντέλου.

  Για τους σκοπούς της εργασίας, ορίζουμε τις εξής παραμέτρους:

  • **Αριθμός χαρακτηριστικών:** Το πλήθος των χαρακτηριστικών που θα χρησιμοποιηθούν στην εκπαίδευση των μοντέλων.

  • **Ακτίνα των clusters ra:** Η παράμετρος που καθορίζει την ακτίνα επιρροής των clusters και κατ΄ επέκταση το πλήθος των κανόνων που θα προκύψουν. Ο καθορισμός των τιμών των παραμέτρων που θα εξεταστούν επιλέγεται ελεύθερα.

- <u>Με βάση τις βέλτιστες τιμές των παραμέτρων που επιλέχθηκαν από το προηγούμενο βήμα :</u> εκπαιδεύουμε ένα τελικό TSK μοντέλο και ελέγχουμε την απόδοσή του στο σύνολο ελέγχου.

ΔΙΑΔΙΚΑΣΙΑ ΜΟΝΤΕΛΟΠΟΙΗΣΗΣ

ΒΗΜΑ 1

Ο διαχωρισμός του συνόλου δεδομένων να γίνει όπως και στο πρώτο κομμάτι, με τα σύνολα εκπαίδευσης-επικύρωσης-ελέγχου να περιλαμβάνουν αντίστοιχα το **60% - 20% - 20%** του συνόλου.

ΒΗΜΑ 2

Να εκτελεστεί αναζήτηση πλέγματος (grid search) και αξιολόγηση μέσω **5-πτυχης διασταυρωμένης επικύρωσης** (5-fold cross validation) για την επιλογή των βέλτιστων τιμών των παραμέτρων. Σε κάθε επανάληψη να αποθηκεύεται το μέσο σφάλμα. Ο διαχωρισμός των δεδομένων να γίνει έτσι ώστε σε κάθε επανάληψη, το **80%** των δεδομένων να χρησιμοποιείται για **εκπαίδευση** και το υπόλοιπο **20%** για **επικύρωση** (ως είσοδοι στη συνάρτηση anfis του MATLAB). Ως μέθοδος **ομαδοποίησης** για τη δημιουργία των IF-THEN κανόνων επιλέγεται ο αλγόριθμος **Subtractive Clustering (SC)** και η επιλογή χαρακτηριστικών μπορεί να εκτελεστεί με έναν από τους εξής αλγορίθμους **(Relief, mRMR, FMI)**. Να εφαρμοστεί **προεπεξεργασία** των δεδομένων αν αυτό κριθεί απαραίτητο.

ΒΗΜΑ 3

Να εκπαιδευτεί το τελικό TSK μοντέλο με τις βέλτιστες τιμές των παραμέτρων και με τις ίδιες προδιαγραφές όπως και προηγουμένως (SC).

****

ΖΗΤΟΥΜΕΝΑ ΠΡΟΒΛΗΜΑΤΟΣ

Μετά το πέρας της διαδικασίας, να σχολιαστούν τα αποτελέσματα όσον αφορά το μέσο σφάλμα σε συνάρτηση με τις τιμές των παραμέτρων. Να δοθούν διαγράμματα τα οποία να απεικονίζουν την καμπύλη αυτού του σφάλματος σε σχέση με τον αριθμό των κανόνων και σε σχέση με τον αριθμό των επιλεχθέντων χαρακτηριστικών. Ποια συμπεράσματα μπορούν να βγουν?

Για τα παρακάτω αποτελέσματα έχει εκτελεστεί το αρχείο **REGRESSION_2_MULTIDIMENTIONAL .m**

Εκτελέστηκαν 2 αλγόριθμοι grid search με τα εξής στοιχεία:

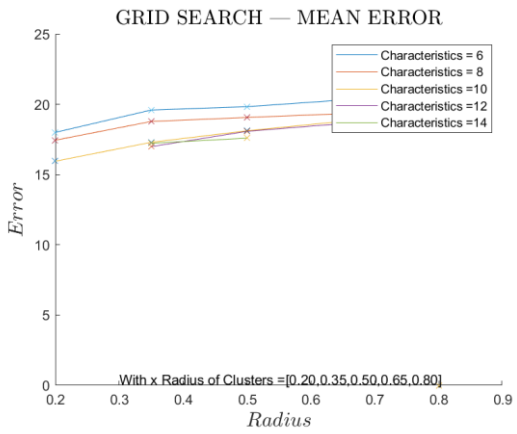
- **Αριθ. Χαρακτηρ.** = [6 8 10 12 14] ,**Ra Επιρ.** = [0.2 0.35 0.5 0.65 0.8]
- **Αριθ. Χαρακτηρ.** = [4 6 8 10 12] ,**Ra Επιρ.** = [0.2 0.35 0.5 0.65 0.8]

Ύστερα από την εκτέλεση των αλγορίθμων grid search επιλέχθηκαν:

- Για τον πρώτο **Αριθ. Χαρακτηρ**. **= 10** και **Ra Επιρ. = 0.2**
- Για το δεύτερο **Αριθ. Χαρακτηρ**. **= 14** και **Ra Επιρ. = 0.8**

Εκτελέσαμε αλγόριθμο grid search με αξιολόγηση μέσω 5-πτυχης διασταυρωμένης επικύρωσης (5-fold cross validation) για την επιλογή των βέλτιστων τιμών των παραμέτρων.

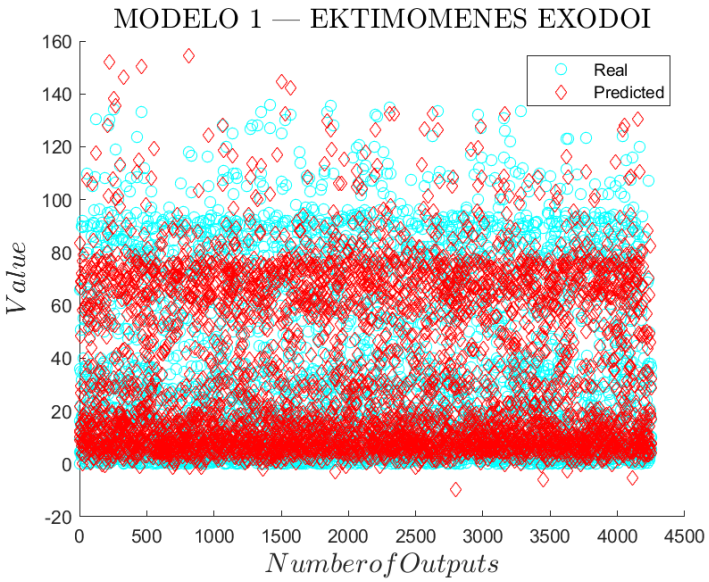
Επίσης έχουμε τα διαγράμματα του μέσου σφάλματος σε συνάρτηση με τις τιμές των παραμέτρων:



Και στα δύο διαγράμματα βλέπουμε πως το σημείο όπου έχει το μικρότερο μέσω σφάλμα είναι εκεί που έχει Ακτίνα επιρροής 0.2 και ο αριθμός χαρακτηριστικών είναι 10 (μωβ-κίτρινη γραμμή).

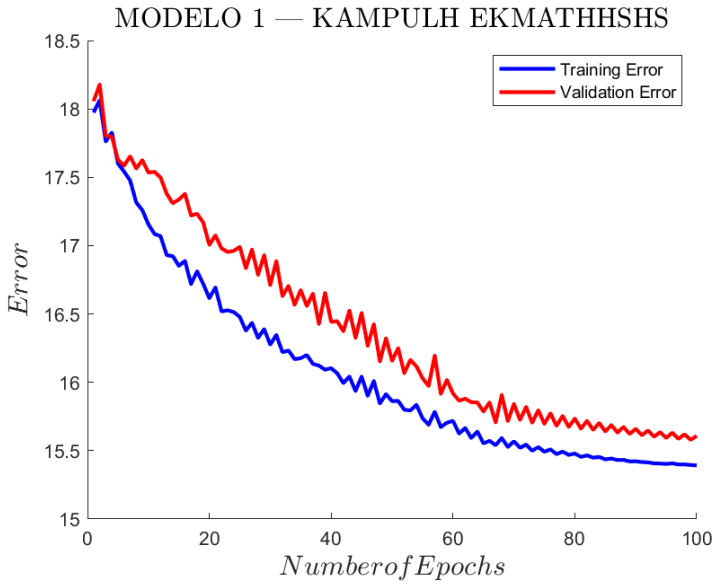
1) ΕΡΩΤΗΜΑ

Διαγράμματα όπου να αποτυπώνονται οι προβλέψεις του τελικού μοντέλου καθώς και οι πραγματικές τιμές.



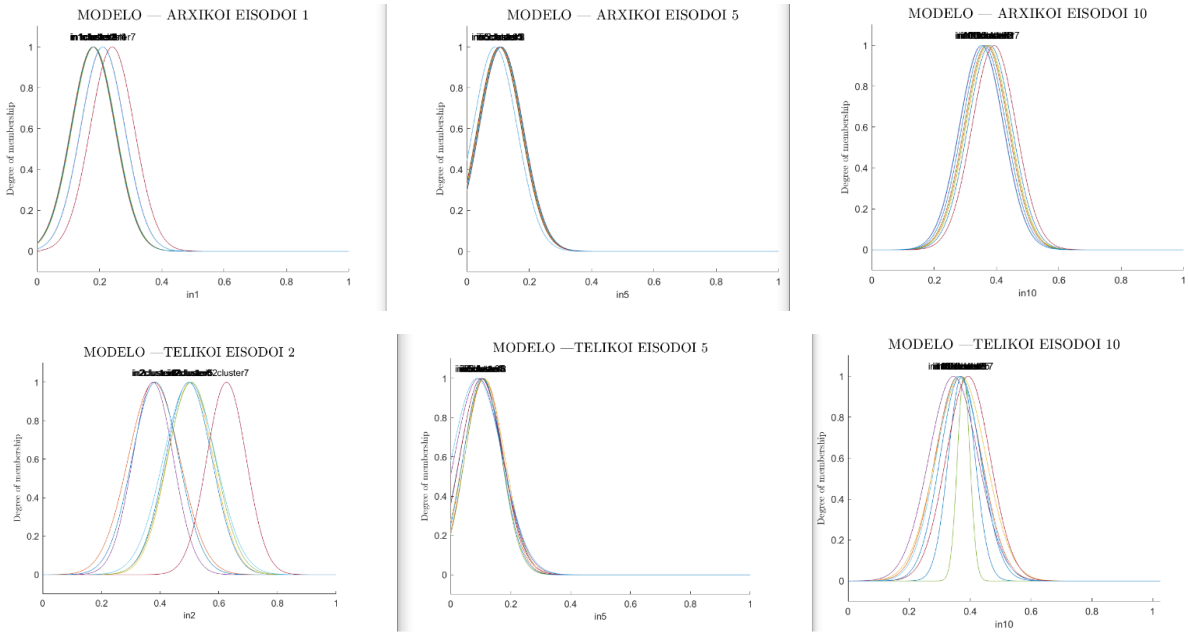
2) ΕΡΩΤΗΜΑ

Διαγράμματα εκμάθησης όπου να απεικονίζεται το σφάλμα συναρτήσει του αριθμού επαναλήψεων.



3) ΕΡΩΤΗΜΑ

Να δοθούν ενδεικτικά μερικά ασαφή σύνολα στην αρχική και τελική τους μορφή.



4) ΕΡΩΤΗΜΑ

Να δοθούν σε ένα πίνακα οι τιμές των δεικτών απόδοσης RMSE, NMSE, NDEI, R2 .

	Βέλτιστο Μοντέλο
R^2	0.8047
RMSE	15.0549
NMSE	0.1953
NDEI	0.4420

ΣΧΟΛΙΑΣΜΟΣ ΑΠΟΤΕΛΕΣΜΑΤΩΝ

Τέλος, να σχολιαστούν τα αποτελέσματα όσον αφορά τα χαρακτηριστικά που επιλέχθηκαν και τον αριθμό IF-THEN κανόνων του ασαφούς συστήματος συμπερασμού. Να γίνει σύγκριση με τον αντίστοιχο αριθμό κανόνων αν για το ίδιο πλήθος χαρακτηριστικών, είχαμε επιλέξει grid partitioning με δύο ή τρία ασαφή σύνολα ανά είσοδο. Ποια είναι τα συμπεράσματα.

Βέλτιστο Μοντέλο

Όπως φαίνεται στο διάγραμμα έχουμε ικανοποιητική προσέγγιση των πραγματικών τιμών από τις εκτιμώμενες τιμές, αυτό μας υπογραμμίζει επίσης ο συντελεστής $R^2 = 0.8047$.
Σύμφωνα με την καμπύλη εκμάθησης που βλέπουμε έχουμε επιτυχή διαδικασία εκπαίδευσης του δικτύου διότι το MSEtrn μειώνεται συναρτήσει των διαδοχικών επαναλήψεων εκπαίδευσης και προσεγγίζει μια σταθερή τιμή . Επίσης το MSEchk μειώνεται συναρτήσει των διαδοχικών επαναλήψεων εκπαίδευσης και προσεγγίζει μια σταθερή τιμή. Έχουμε μικρή απόσταση(διαφορά) του TrainingError και το ValidationError . Σε αυτό το μοντέλο δεν εμφανίζεται το φαινόμενο τις υπερεκπαίδευσης καθώς χρησιμοποιήσαμε το σύνολο αξιολογήσης Dval το οποίο αποτελεί ένα εργαλείο

αποτελεσματικής εκπαίδευσης (υψηλή ακρίβεια προσέγγισης) και δημιουργίας μοντέλων με αυξημένες ιδιότητες γενίκευσης. (χρησιμοποιήσαμε την τεχνική Cross Validation) .

Οι IF-THEN κανόνων του ασαφούς συστήματος συμπερασμού που χρησιμοποιήθηκαν είναι 9 σε αριθμό. Αν είχαμε επιλέξει grid partitioning με δύο ή τρία ασαφή σύνολα ανά είσοδο τότε θα είχαμε **2^10** (1024) ή **3^10** (59049) κανόνες , όπου θα οδηγούσαν σε ένα πολύπλοκο μοντέλο που θα ήθελα τεράστια υπολογιστική ισχύει για να μας δώσει αποτελέσματα. Άρα στην ουσία το πρόβλημα θα ήταν σχεδόν αδύνατο να λυθεί.

****