

ΕΡΓΑΣΙΑ ΣΤΟ ΜΑΘΗΜΑ ΑΣΑΦΗ ΣΥΣΤΗΜΑΤΑ

Επίλυση προβλήματος παλινδρόμησης με χρήση μοντέλων TSK

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 επίδεξη τρόπων ανάλυσης και ερμηνείας των αποτελεσμάτων. Το δεύτερο, πολυπλοκότερο σύνολο δεδομένων θα χρησιμοποιηθεί για μια πληρέστερη διαδικασία μοντελοποίησης, η οποία θα περιλαμβάνει μεταξύ άλλων προεπεξεργαστικά βήματα όπως επιλογή χαρακτηριστικών (feature selection), καθώς και μεθόδους βελτιστοποίησης των μοντέλων μέσω της διασταυρωμένης επικύρωσης (cross validation).

Λίγα λόγια για το πρόβλημα της παλινδρόμησης: Το πρόβλημα της παλινδρόμησης αποτελεί ένα υποσύνολο του γενικότερου ζητήματος της αναγνώρισης προτύπων, και ασχολείται με την εκτίμησης μιας άγνωστης συνάρτησης με βάση κάποια διαθέσιμα δεδομένα. Συγκεκριμένα, ας υποθέσουμε ότι έχουμε ένα σύνολο δειγμάτων $X = \{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N\}$, όπου N είναι το πλήθος των δειγμάτων και $\vec{x}_i \in R^M$, με M το πλήθος των χαρακτηριστικών (attributes), ένα σύνολο εξόδων $Y = \{y_1, y_2, \dots, y_N\}$, $y_i \in R$, και ότι υπάρχει μια πραγματική συνάρτηση $f(\vec{x}): R^M \rightarrow R$ η οποία απεικονίζει τα στοιχεία του συνόλου X στα στοιχεία του συνόλου Y , θα έχουμε δηλαδή $y_i = f(\vec{x}_i)$. Στόχος μας είναι, με χρήση των διαθέσιμων δεδομένων, να κατασκευάσουμε μια συνάρτηση $\hat{f}()$ η οποία να αποτελεί μια όσο το δυνατόν καλύτερη προσέγγιση της πραγματικής συνάρτησης $f()$.

1) Μια πρώτη εφαρμογή σε UCI dataset: Στην πρώτη φάση της εργασίας, επιλέγεται από το UCI repository το Cobined Cycle Power Plant (CCPP) dataset, το οποίο περιλαμβάνει 9568 δείγματα (instances) και 4 χαρακτηριστικά (features). Τα δεδομένα συλλέχθηκαν από ένα εργοστάσιο παραγωγής ηλεκτρικής ενέργειας κατά τη διάρκεια 6 ετών και χαρακτηρίζονται από τις εξής τέσσερις μεταβλητές:

- Μέση ωριαία θερμοκρασία (T)
- Μέση ωριαία πίεση (AP)
- Μέση ωριαία σχετική υγρασία (RH)
- Μέση ωριαία exhaust vacuum (V)

Η μεταβλητή στόχος, η μεταβλητή δηλαδή της οποίας την τιμή θέλουμε να προβλέψουμε χρησιμοποιώντας τις παραπάνω τέσσερις μεταβλητές χαρακτηριστικών, είναι η ενεργειακή απόδοση του σταθμού (energy output). 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 $\{D_{trn}, D_{val}, D_{chk}\}$, από τα οποία το πρώτο θα χρησιμοποιηθεί για εκπαίδευση, το δεύτερο για επικύρωση και αποφυγή του φαινομένου

υπερεκπαίδευσης και το τελευταίο για τον έλεγχο της απόδοσης του τελικού μας μοντέλου. Προτείνεται να χρησιμοποιηθεί το 60% του συνόλου των δειγμάτων για το υποσύνολο εκπαίδευσης και από 20% του συνόλου των δειγμάτων για κάθε ένα από τα δύο εναπομείναντα υποσύνολα.

- 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. Δίνεται ο παρακάτω πίνακας για περαιτέρω επεξήγηση:

	Πλήθος συναρτήσεων συμμετοχής	Μορφή εξόδου
<i>TSK_model_1</i>	2	Singleton
<i>TSK_model_2</i>	3	Singleton
<i>TSK_model_3</i>	2	Polynomial
<i>TSK_model_4</i>	3	Polynomial

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 τέτοιο τρόπο ώστε τα διαδοχικά ασαφή σύνολα να παρουσιάζουν σε κάθε είσοδο, βαθμό επικάλυψης περίπου 0.5.

- 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: το μέσο τετραγωνικό σφάλμα μεταξύ της εξόδου του μοντέλου και της πραγματικής εξόδου, όπου θ το διάνυσμα παραμέτρων του τελικού μοντέλου που έχει επιλεχθεί.

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

Εναλλακτικά, μπορεί να χρησιμοποιηθεί ο δείκτης RMSE, ο οποίος είναι απλά η τετραγωνική ρίζα του δείκτη MSE: $RMSE = \sqrt{MSE(\vec{\theta})}$.

2. Ο συντελεστής προσδιορισμού R^2 , ο οποίος μας δίνει πληροφορία για το ποσοστό της διακύμανσης της πραγματικής το οποίο “εξηγείται” από το μοντέλο μας. Ο υπολογισμός του συντελεστή προσδιορισμού μπορεί να γίνει με τον εξής τρόπο:

$$SS_{res} = \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

$$SS_{tot} = \sum_{i=1}^N (y_i - \bar{y})^2 \quad ,$$

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

- όπου y είναι η πραγματική έξοδος του συστήματος, \hat{y} η εκτίμηση που παράγει το μοντέλο μας και \bar{y} η μέση τιμή της πραγματικής εξόδου. Ο συντελεστής προσδιορισμού συνήθως λαμβάνει τιμές στο διάστημα $[0,1]$ με τη μονάδα να αντιστοιχεί σε τέλεια εκτίμηση και το μηδέν σε μοντέλο το οποίο μονίμως παράγει ως εκτίμηση την μέση τιμή της πραγματικής εξόδου.
3. Τέλος, για πιο εύρωστη αξιολόγηση, θα μπορούσαν να χρησιμοποιηθούν οι δείκτες NMSE και NDEI οι οποίοι υπολογίζονται ως εξής:

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

$$NDEI = \sqrt{NMSE} = \frac{\sigma_e}{\sigma_x}$$

- Ζητούμενα του προβλήματος: 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 περιγράφηκαν παραπάνω. Ως τελικό μοντέλο να επιλέγεται πάντα εκείνο το οποίο αντιστοιχεί στο μικρότερο σφάλμα στο σύνολο επικύρωσης. Για τις τέσσερις περιπτώσεις εκπαίδευσης:
 1. Να δώσετε τα αντίστοιχα διαγράμματα στα οποία να απεικονίζονται οι τελικές μορφές των ασαφών συνόλων που προέκυψαν μέσω της διαδικασίας εκπαίδευσης.
 2. Να δοθούν τα διαγράμματα μάθησης (learning curves) όπου να απεικονίζεται το σφάλμα του μοντέλου συναρτήσει του αριθμού των επαναλήψεων (iterations).
 3. Να δοθούν τα διαγράμματα όπου να αποτυπώνονται τα σφάλματα πρόβλεψης.
 4. Τέλος, να παρουσιαστούν σε μορφή πίνακα οι τιμές των δεικτών απόδοσης $RMSE, NMSE, NDEI, R^2$.

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.

2) Εφαρμογή σε dataset με υψηλή διαστασιμότητα: 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. Ένα 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.

Σύμφωνα με τα παραπάνω, μια αρκετά προφανής προσέγγιση θα ήταν να προσπαθήσουμε να μειώσουμε ταυτόχρονα και το πλήθος των εισόδων και το πλήθος των κανόνων. Οι δύο τεχνικές που παρουσιάζονται παρακάτω στοχεύουν σε αυτόν ακριβώς το διττό σκοπό.

1. Επιλογή χαρακτηριστικών (feature selection): Η επιλογή χαρακτηριστικών αποτελεί μια από τις πιο διαδεδομένες και πλούσιες οικογένειες τεχνικών για τη μείωση της διαστασιμότητας σε ένα πρόβλημα μοντελοποίησης. Η βασική ιδέα πίσω από αυτές τις τεχνικές είναι ότι στη συντριπτική πλειοψηφία των περιπτώσεων, ένας σημαντικός αριθμός χαρακτηριστικών/εισόδων των δειγμάτων ενός συνόλου δεδομένων είναι πλεονάζοντα, επομένως θα μπορούσαν να παραλειφθούν από τη διαδικασία μοντελοποίησης. Αυτό μας επιτρέπει να μειώσουμε από τη μία τη διαστασιμότητα του προβλήματος, καθιστώντας το ευκολότερα διαχειρίσιμο, και από την άλλη, να απλοποιήσουμε το τελικό μοντέλο, έτσι ώστε να είναι ευκολότερα ερμηνεύσιμο από τους εκάστοτε ερευνητές/χρήστες. Για τους σκοπούς της συγκεκριμένης εργασίας, ως μέθοδος επιλογής χαρακτηριστικών επιλέγεται ο αλγόριθμος *Relief*. Περισσότερες λεπτομέρειες για τη λογική του αλγορίθμου θα βρεθούν στο σχετικό paper που έχει ανέβει στην ενότητα Υλικό Μαθήματος (ethmmy).
2. Διαμέριση διασκορπισμού (scatter partitioning): Κατά τη διαδικασία αρχικοποίησης ενός TSK μοντέλου, ένα από τα βασικά βήματα είναι ο διαχωρισμός του χώρου εισόδου και η δημιουργία των αρχικών ασαφών συνόλων. Η απλούστερη μέθοδος προς το σκοπό αυτό είναι το grid partitioning, όπως αναφέρθηκε όμως παραπάνω, η συγκεκριμένη μέθοδος οδηγεί σε εκθετικά αυξανόμενο πλήθος κανόνων σε συνάρτηση με το πλήθος των εισόδων. Μια εναλλακτική επιλογή είναι η χρήση μεθόδων ομαδοποίησης για το διαχωρισμό του χώρου εισόδου και η αρχικοποίηση των ασαφών συνόλων πάνω στις ομάδες που προέκυψαν. Με τον τρόπο αυτό, το πλήθος των κανόνων παύει να εξαρτάται από το πλήθος των εισόδων και εξαρτάται πλέον αποκλειστικά από τον αριθμό των clusters. Επιπλέον, οι κανόνες του ασαφούς συστήματος είναι διασκορπισμένοι (scattered) στο χώρο εισόδου και δε δημιουργούν ένα πλέγμα όπως στην περίπτωση του grid partitioning. Δύο δημοφιλείς επιλογές για την ομαδοποίηση των δεδομένων και τη μετέπειτα δημιουργία των κανόνων είναι η αφαιρετική ομαδοποίηση (subtractive clustering) και ο αλγόριθμος Fuzzy C-means (FMC).

Το dataset που θα επιλεγεί για την επίδειξη των παραπάνω μεθόδων είναι το Bank dataset από το delve repository. Το συγκεκριμένο dataset, περιλαμβάνει 8192 δείγματα, καθένα από τα οποία περιγράφεται από 33 μεταβλητές/χαρακτηριστικά. Πρόκειται για ένα τεχνητό dataset το οποίο παράχθηκε από έναν simulator με σκοπό τη χρήση του σε προβλήματα regression. Η προσομοίωση είχε να κάνει με το μέγεθος των ουρών αναμονής σε μια τράπεζα και ο σκοπός είναι να εκτιμηθεί η πιθανότητα ένας πελάτης να αποχωρήσει λόγω της μεγάλης αναμονής. Είναι φανερό ότι το μέγεθος του dataset καθιστά απαγορευτική μια απλή εφαρμογή ενός TSK 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 32 μεταβλητές/predictors, διαμερίζαμε το χώρο εισόδου κάθε μεταβλητής με δύο ασαφή

σύνολα, θα καταλήγαμε με 2^{23} κανόνες). Ο στόχος αυτός θα επιτευχθεί μέσω της επιλογής χαρακτηριστικών και της χρήσης διαμέρισης διασκορπισμού, όπως συζητήθηκε παραπάνω. 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:

1. Διαχωρισμός σε σύνολα εκπαίδευσης- επικύρωσης – ελέγχου: Όπως και στο πρώτο κομμάτι της εργασίας, είναι απαραίτητος ο διαχωρισμός του συνόλου δεδομένων σε τρία υποσύνολα $\{D_{trn}, D_{val}, D_{chk}\}$, το ένα από τα οποία θα χρησιμοποιηθεί για εκπαίδευση και το δεύτερο για έλεγχο της απόδοσης.
2. 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 ελέγχουμε 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. Η όλη διαδικασία γίνεται ευκολότερα κατανοητή με το παρακάτω κομμάτι ψευδοκώδικα:

```
1 Define sets of model parameter values to evaluate
2 for each parameter set do
3     for each resampling iteration do
4         Hold-out specific samples
5         [Optional] Pre-process the data
6         Fit the model on the remainder
7         Predict the hold-out samples
8     end
9     Calculate the average performance across hold-out predictions
10 end
11 Determine the optimal parameter set
12 Fit the final model to all the training data using the optimal parameter set
```

Για τους σκοπούς αυτής της εργασίας, ορίζουμε τις εξής παραμέτρους:

- Αριθμός χαρακτηριστικών: Οι τιμές των χαρακτηριστικών που θα χρησιμοποιηθούν δίνονται από το σύνολο $NF=\{3,9,15,21\}$.
- Αριθμός κανόνων: Ο αριθμός κανόνων (αριθμός των clusters), θα λαμβάνει τιμές από το σύνολο $NR=\{5,10,15,20,25\}$.

3. Με βάση τις βέλτιστες τιμές των παραμέτρων που επιλέχθηκαν από το προηγούμενο βήμα, εκπαιδεύουμε ένα τελικό TSK μοντέλο και ελέγχουμε την απόδοσή του στο σύνολο ελέγχου.

Τα παραπάνω βήματα συνοψίζουν πλήρως τη διαδικασία μοντελοποίησης που θα ακολουθηθεί. Σημειώνεται ότι τα σύνολα παραμέτρων έτσι όπως παρουσιάζονται παραπάνω είναι προαιρετικά, και μπορεί κανείς να αντικαταστήσει τις τιμές τους, ειδικά αν η διαδικασία της διασταυρωμένης επικύρωσης αποδειχθεί ιδιαίτερα χρονοβόρα. Ζητούνται τα εξής:

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.
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). Μια συνάρτηση που μπορεί να βοηθήσει σε αυτό το έργο είναι η `cvpartition()`. Ως μέθοδος ομαδοποίησης για τη δημιουργία των IF-THEN κανόνων επιλέγεται ο αλγόριθμος Subtractive Clustering (SC – `genfis2`). Να εφαρμοστεί προεπεξεργασία των δεδομένων αν αυτό κριθεί απαραίτητο. 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 επιλεγμένων χαρακτηριστικών. Ποιά συμπεράσματα μπορούν να βγουν;
3. Να εκπαιδευτεί το τελικό TSK μοντέλο με τις βέλτιστες τιμές των παραμέτρων και με τις ίδιες προδιαγραφές όπως και προηγουμένως (SC). Να δοθούν τα εξής διαγράμματα:
 - Διαγράμματα όπου να αποτυπώνονται οι προβλέψεις του τελικού μοντέλου καθώς και οι πραγματικές τιμές.
 - Διαγράμματα εκμάθησης όπου να απεικονίζεται το σφάλμα συναρτήσεως του αριθμού επαναλήψεων.
 - Να δοθούν ενδεικτικά μερικά ασαφή σύνολα στην αρχική και τελική τους μορφή.
 - Να δοθούν σε ένα πίνακα οι τιμές των δεικτών απόδοσης $RMSE, NMSE, NDEI, R^2$.
 - 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. Ποιά είναι τα συμπεράσματα;