

Υπολογιστική Νοημοσύνη

Αναφορά στα πλαίσια της 4ης εργασίας

Classification



Τσαλαγεώργος Βασίλειος

A.E.M. 8253

Εαρινό εξάμηνο 2021

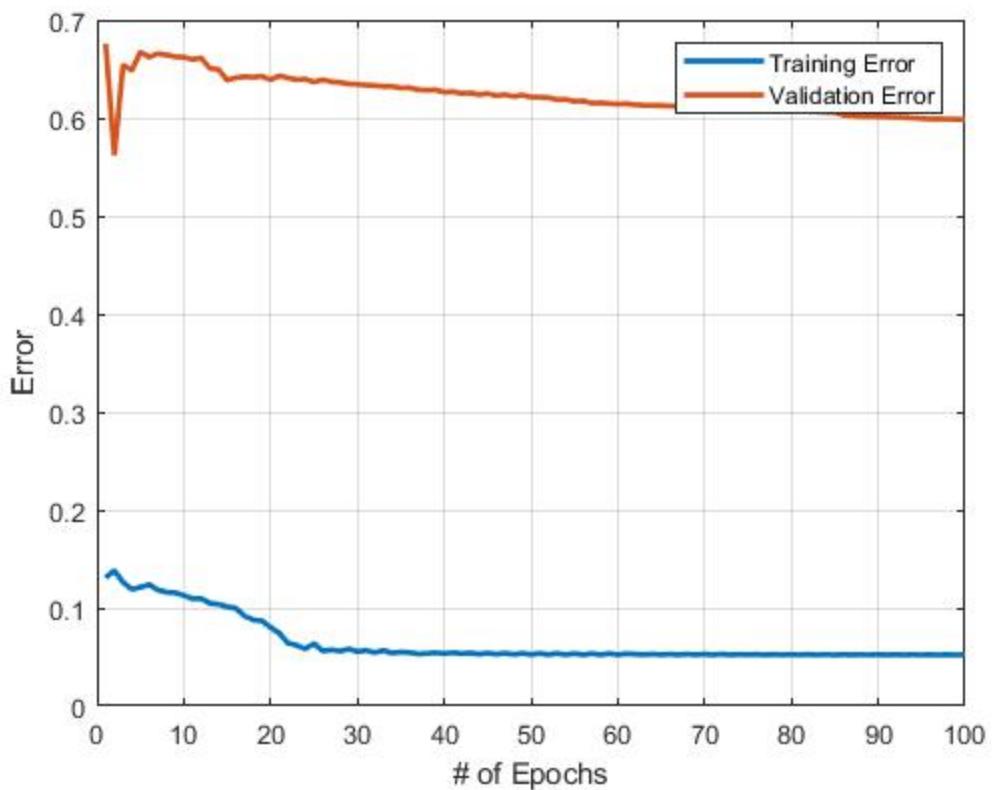
Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης

Τμήμα Ηλεκτρολόγων Μηχανικών & Μηχανικών Υπολογιστών

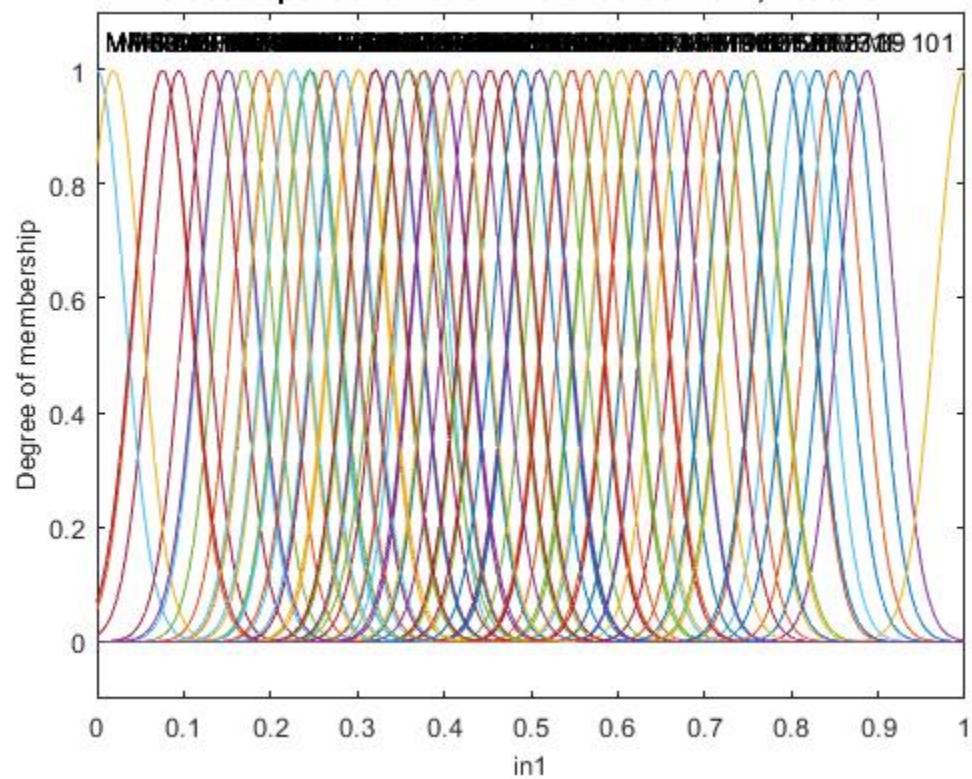
Εφαρμογή σε απλό dataset

Το πρώτο μέρος της εργασίας υλοποιείται σε κώδικα MATLAB και αποθηκεύεται στο αρχείο main.m. Σε πρώτη φάση, γίνεται ο διαχωρισμός των δεδομένων σύμφωνα με την εκφώνηση, σε τρία υποσύνολα εκπαίδευσης (trnData), επικύρωσης (chkData), και ελέγχου (tstData), σε ποσοστά 60%, 20% και 20% αντίστοιχα. Εν συνεχεία, ζητείται για τις δύο περιπτώσεις που θα εξεταστούν (class dependent και class independent), η παράμετρος που καθορίζει το μέγεθος των clusters και τον αριθμό των κανόνων να λάβει δύο ακραίες τιμές, ώστε να εμφανιστεί σημαντική διαφορά στον αριθμό των κανόνων ανάμεσα στα δύο μοντέλα (δύο μοντέλα για κάθε περίπτωση, ένα για κάθε ακραία τιμή της παραμέτρου, τέσσερα συνολικά). Η εν λόγω παράμετρος είναι η ακτίνα των clusters, εφόσον αυτή καθορίζει την ακτίνα επιρροής τους και τον αριθμό των κανόνων που θα προκύψουν. Επιλέγονται το 0.1 και το 0.9 ως οι δύο ακραίες τιμές, καθώς η ακτίνα λαμβάνει τιμές στο [0,1]. Στη συνέχεια, ξεκινά η εκπαίδευση των δύο μοντέλων (class dependent και class independent), σε δύο επαναλήψεις, μία για κάθε τιμή της ακτίνας. Δημιουργούμε επίσης τα διαγράμματα των συναρτήσεων συμμετοχής του κάθε μοντέλου για κάθε χαρακτηριστικό, το διάγραμμα μάθησης του κάθε μοντέλου, όπου απεικονίζεται το σφάλμα συναρτήσει του αριθμού των επαναλήψεων, καθώς επίσης και τον πίνακα σφαλμάτων ταξινόμησης, τον οποίο χρησιμοποιούμε κάθε φορά ώστε να εξάγουμε για κάθε μοντέλο τις τιμές των δεικτών απόδοσης που μας ενδιαφέρουν.

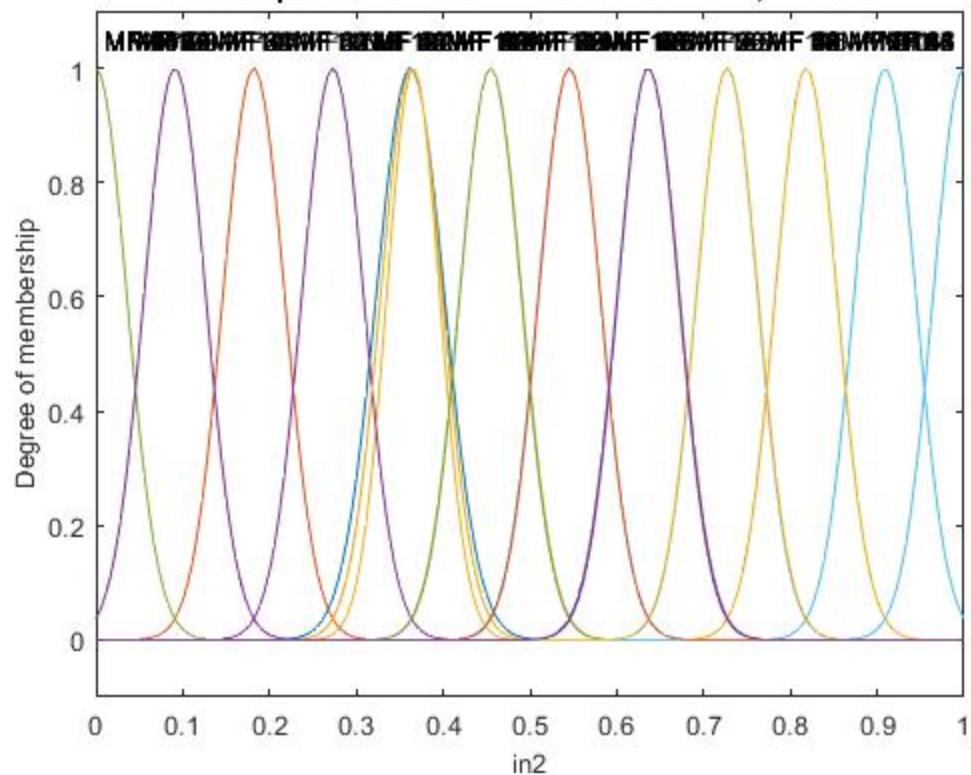
Παρατίθενται εδώ όλα τα παραπάνω σύμφωνα με τα ζητούμενα του προβλήματος, πρώτα για τα class dependent μοντέλα με ακτίνα 0.1 το πρώτο και 0.9 το δεύτερο, και στη συνέχεια για τα class independent μοντέλα με ακτίνες 0.1 και 0.9 αντίστοιχα.



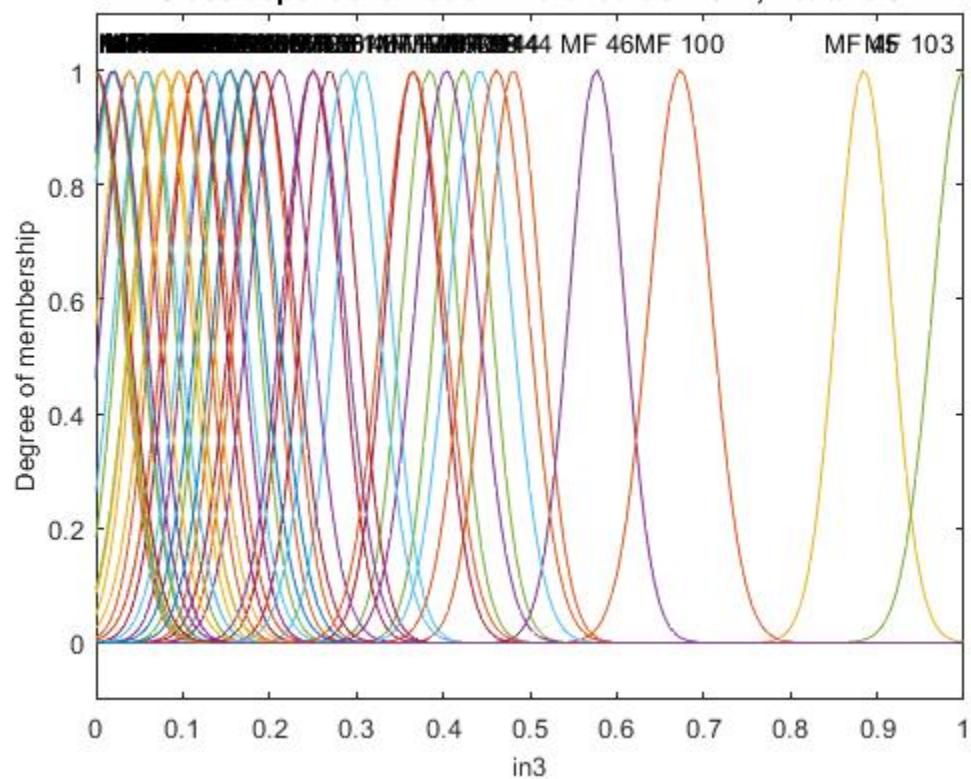
Class dependent model where radius = 0.1 , Feature 1

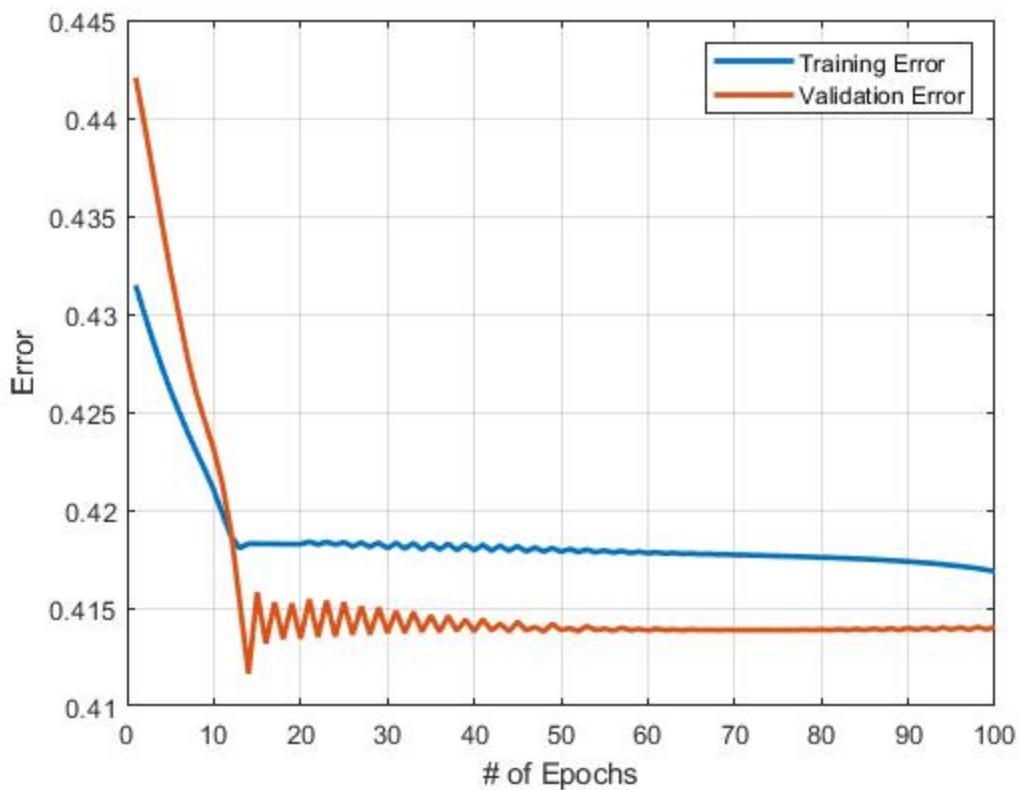


Class dependent model where radius = 0.1 , Feature 2

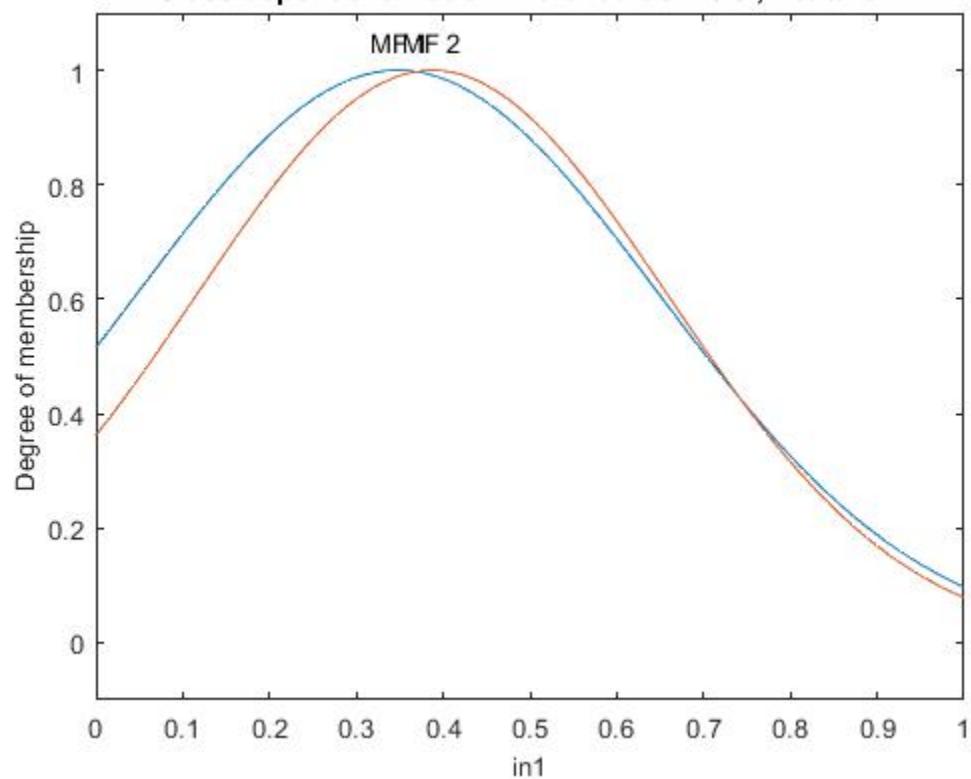


Class dependent model where radius = 0.1 , Feature 3

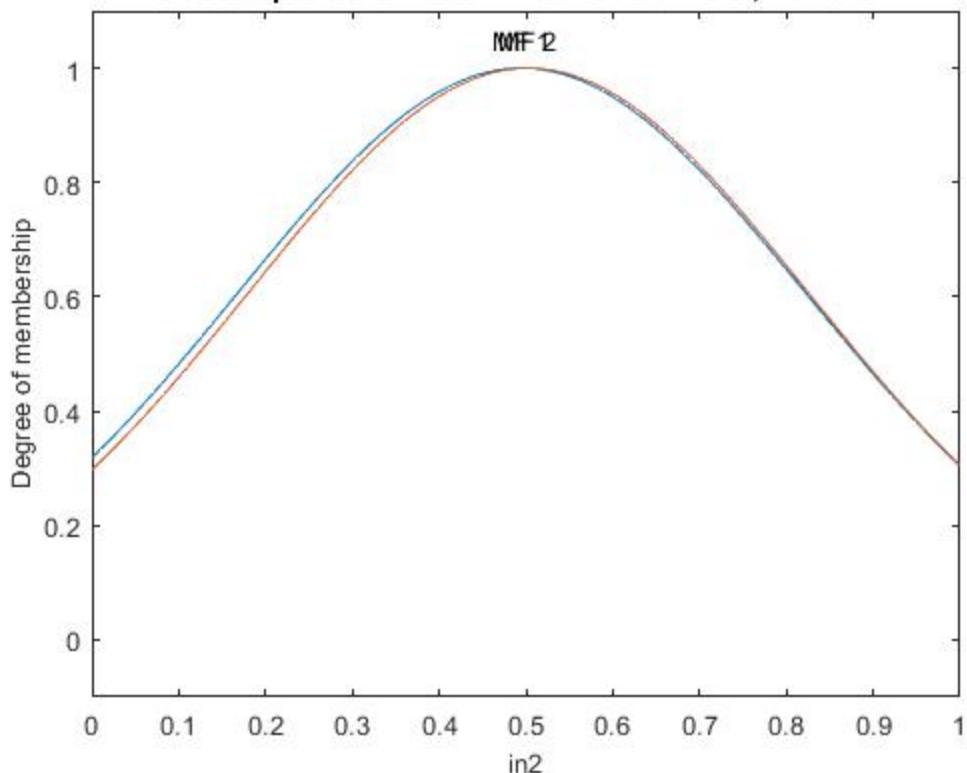




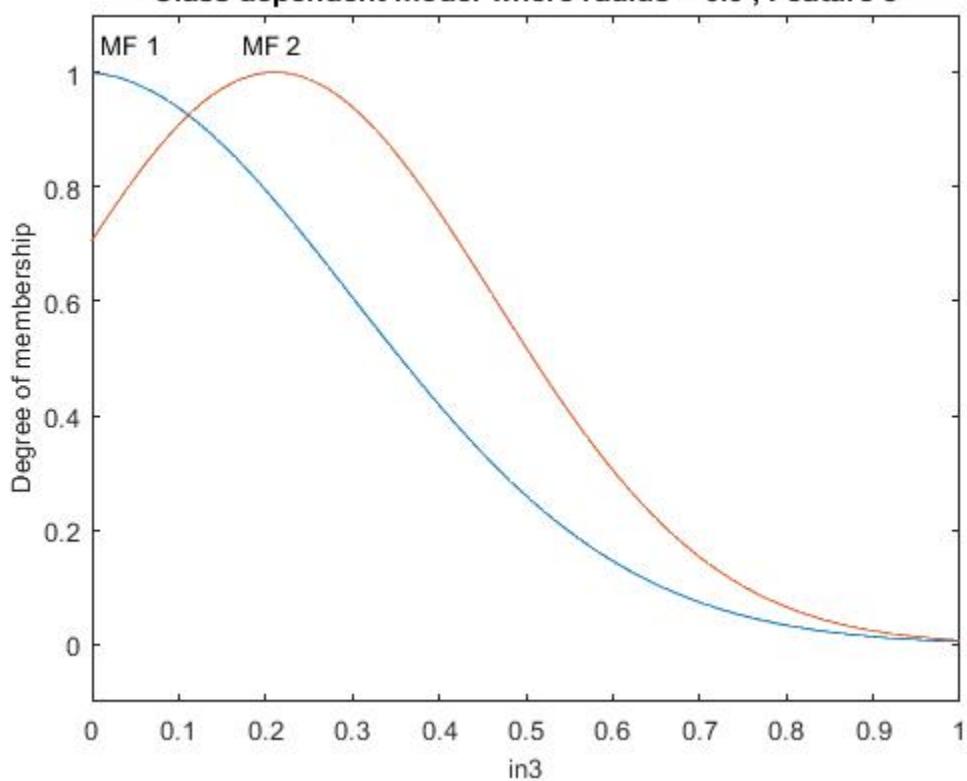
Class dependent model where radius = 0.9 , Feature 1

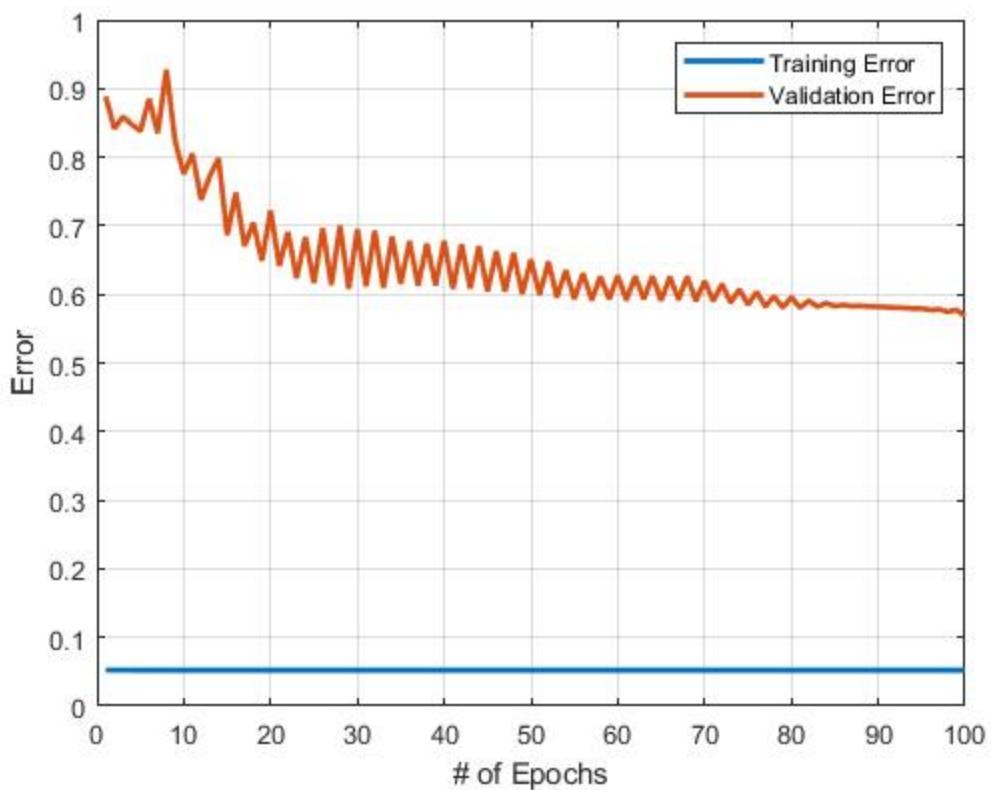


Class dependent model where radius = 0.9 , Feature 2

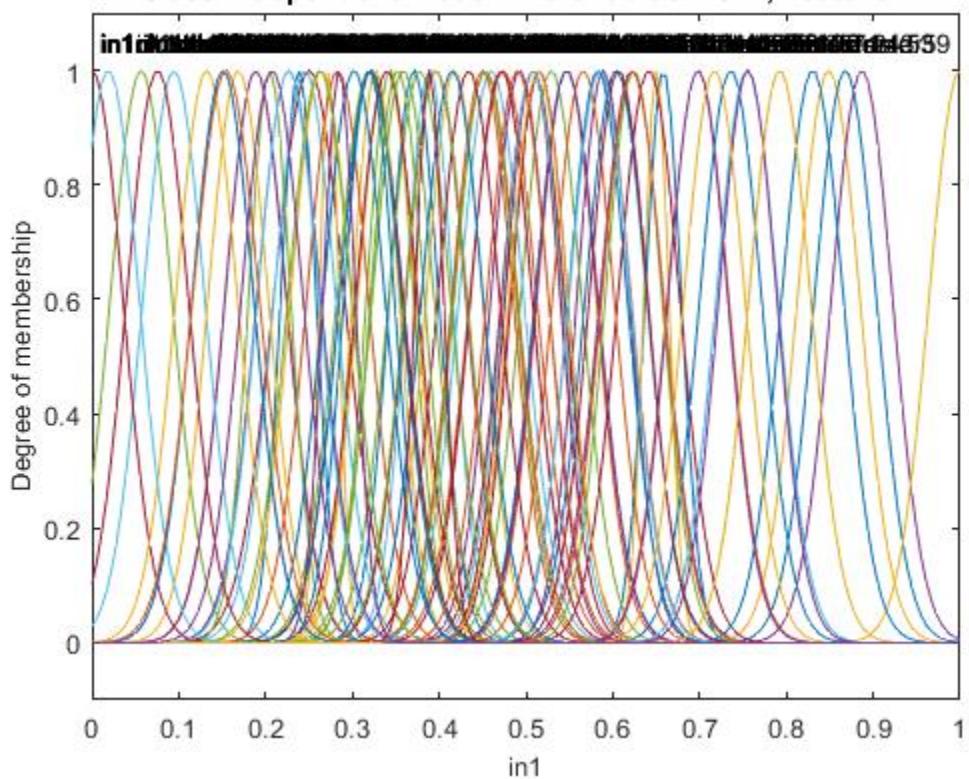


Class dependent model where radius = 0.9 , Feature 3

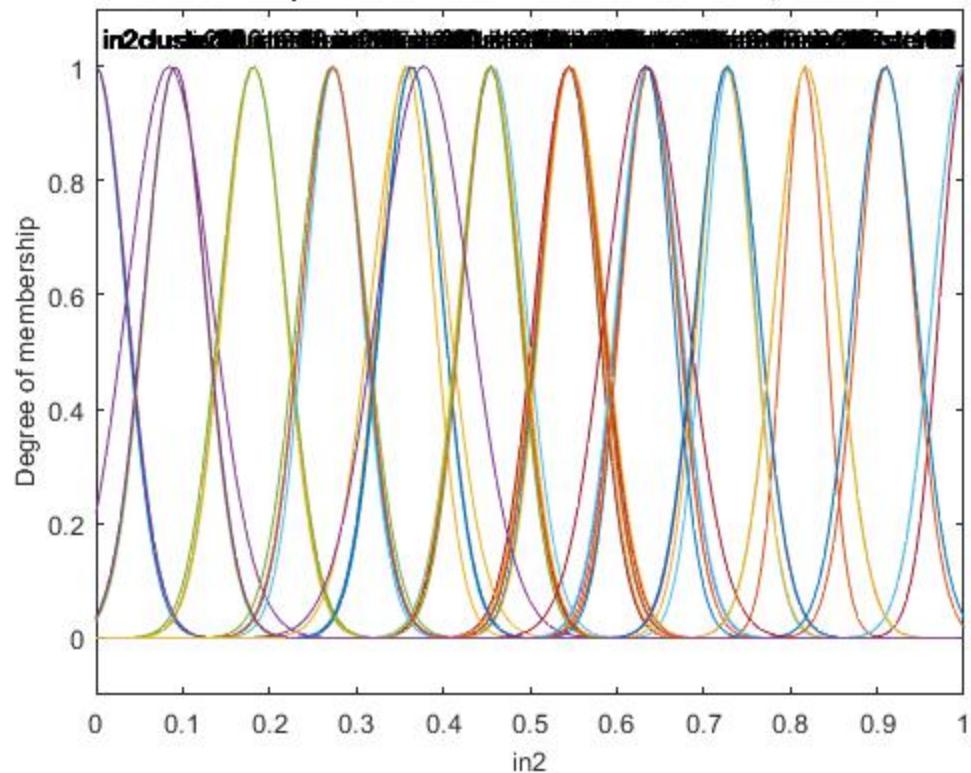




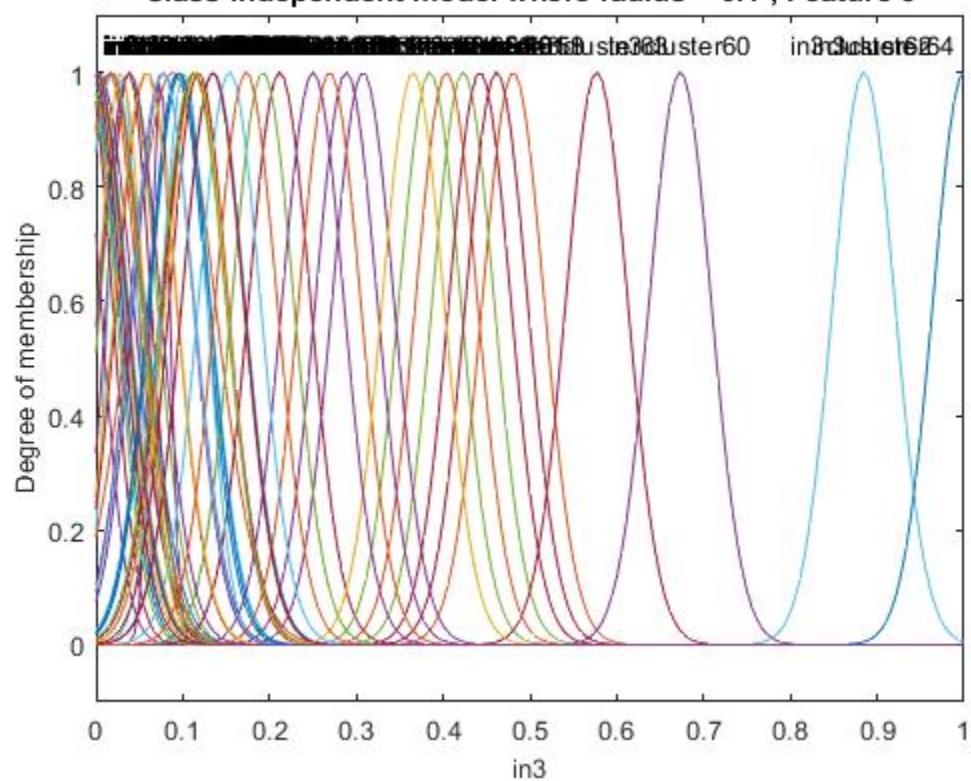
Class independent model where radius = 0.1 , Feature 1

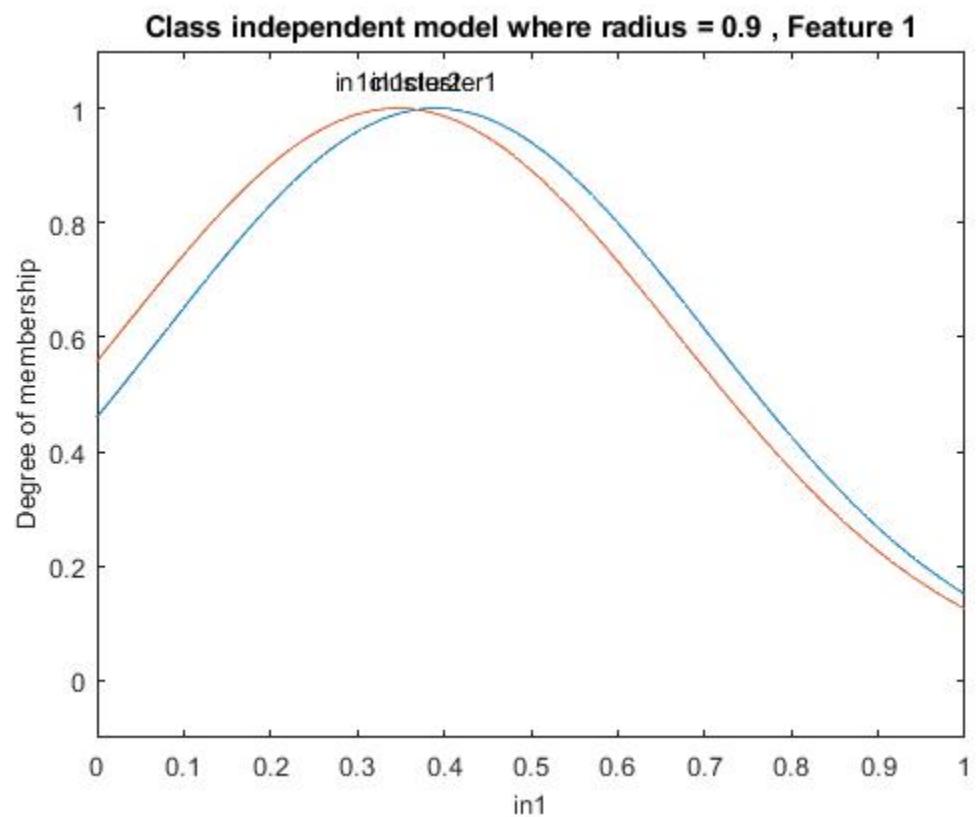
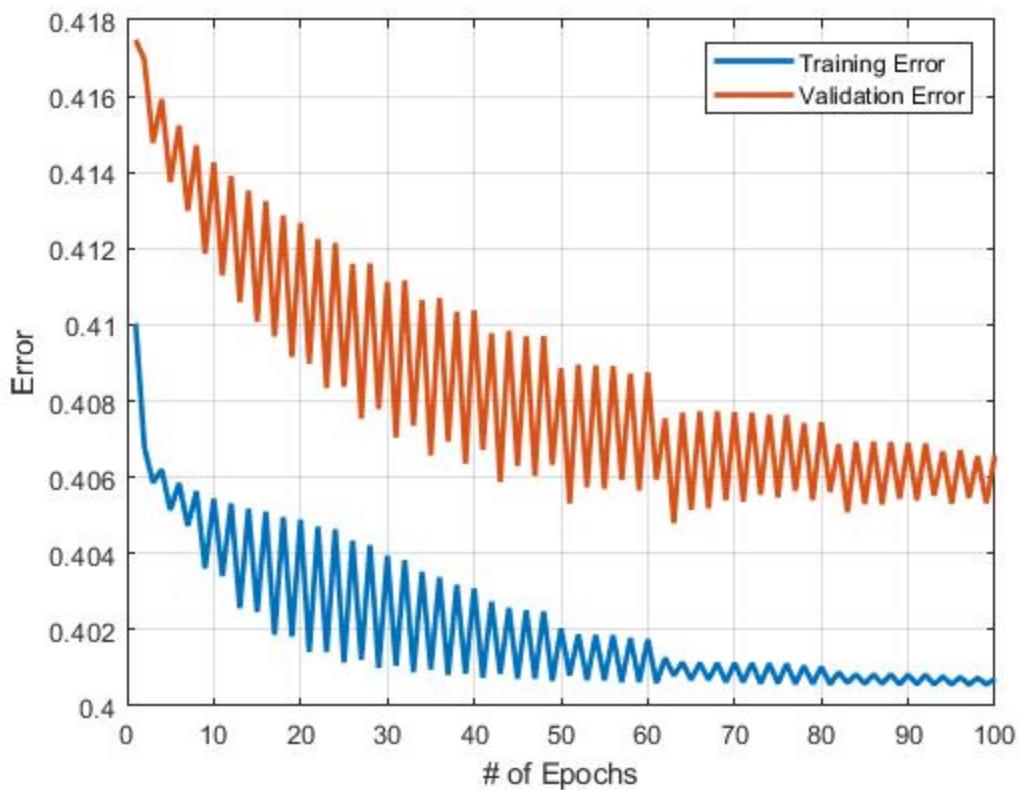


Class independent model where radius = 0.1 , Feature 2

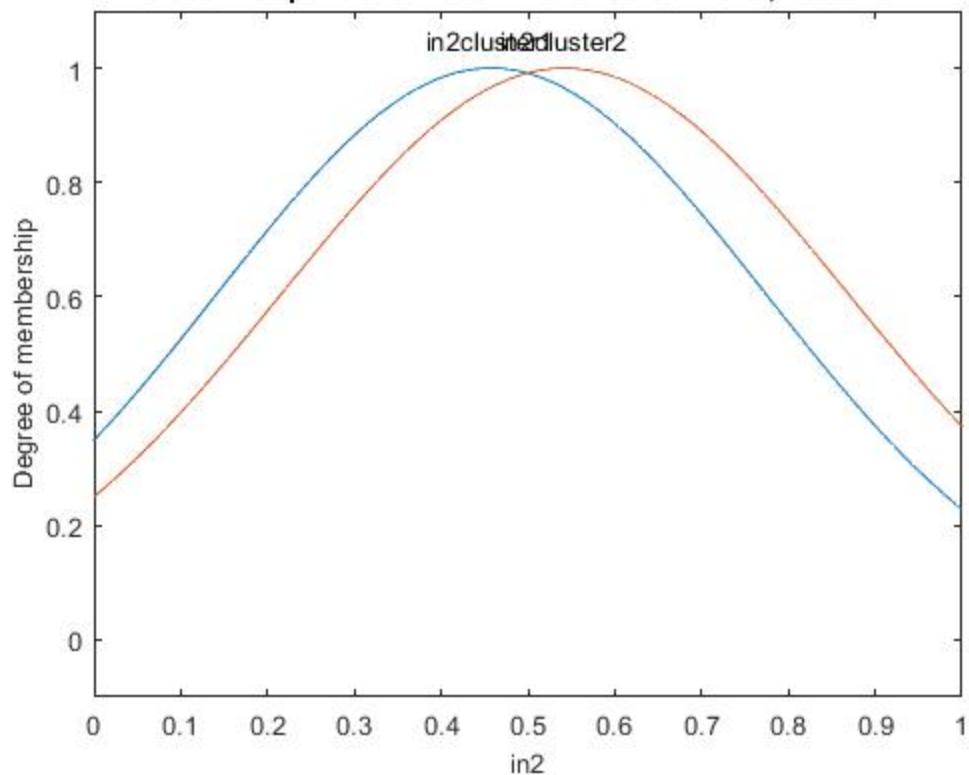


Class independent model where radius = 0.1 , Feature 3

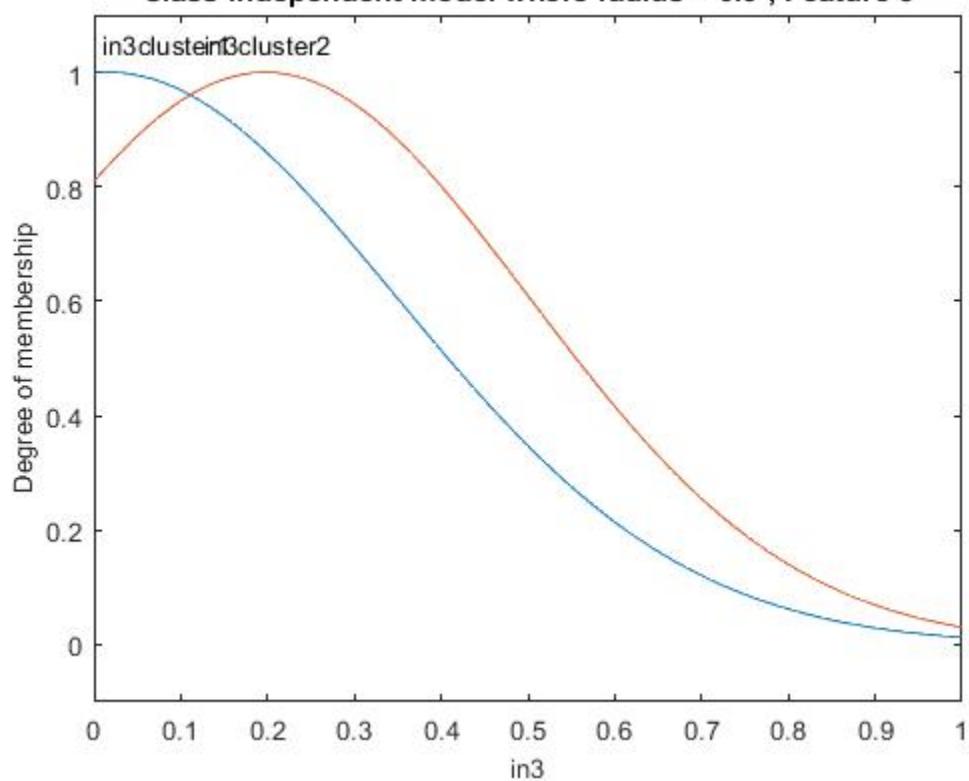




Class independent model where radius = 0.9 , Feature 2



Class independent model where radius = 0.9 , Feature 3



Ακολουθούν οι πίνακες σφαλμάτων και οι τιμές των δεικτών απόδοσης

Πίνακας σφαλμάτων ταξινόμησης για το class dependent μοντέλο με ακτίνα 0.1

40	7
13	1

Τιμές των δεικτών απόδοσης για το class dependent μοντέλο με ακτίνα 0.1

OA	0.6721
PA for class 1	0.8511
PA for class 2	0.0714
UA for class 1	0.7547
UA for class 2	0.1250
K	-0.0912

Αριθμός κανόνων του class dependent μοντέλου με ακτίνα 0.1: 130

Πίνακας σφαλμάτων ταξινόμησης για το class dependent μοντέλο με ακτίνα 0.9

46	1
12	2

Τιμές των δεικτών απόδοσης για το class dependent μοντέλο με ακτίνα 0.9

OA	0.7869
PA for class 1	0.9787
PA for class 2	0.1429
UA for class 1	0.7931
UA for class 2	0.6667
K	0.1679

Αριθμός κανόνων του class dependent μοντέλου με ακτίνα 0.9: 2

Πίνακας σφαλμάτων ταξινόμησης για το class independent μοντέλο με ακτίνα 0.1

36	11
13	1

Τιμές των δεικτών απόδοσης για το class independent μοντέλο με ακτίνα 0.1

OA	0.6066
PA for class 1	0.7660
PA for class 2	0.0714
UA for class 1	0.7347
UA for class 2	0.0833
K	-0.1712

Αριθμός κανόνων του class independent μοντέλου με ακτίνα 0.1: 123

Πίνακας σφαλμάτων ταξινόμησης για το class independent μοντέλο με ακτίνα 0.9

44	3
14	0

Τιμές των δεικτών απόδοσης για το class independent μοντέλο με ακτίνα 0.9

OA	0.7213
PA for class 1	0.9362
PA for class 2	0
UA for class 1	0.7586
UA for class 2	0
K	-0.0881

Αριθμός κανόνων του class independent μοντέλου με ακτίνα 0.9: 2

Βλέπουμε στα παραπάνω, πως το class dependent μοντέλο με ακτίνα 0.9, πετυχαίνει το καλύτερο overall accuracy, τα καλύτερα producer's και user's accuracies και για τις δύο κλάσεις, και το καλύτερο (και μόνο θετικό) K. Συνεπώς, πρόκειται αδιαμφισβήτητα για το βέλτιστο μοντέλο από τα τέσσερα που εκπαιδεύτηκαν συνολικά. Παρατηρούμε επίσης, πως όσο μικραίνουμε την τιμή της ακτίνας, τόσο αυξάνεται ο αριθμός των κανόνων, και συγκρίνοντας τα dependent μοντέλα μεταξύ τους, καθώς και τα independent αντίστοιχα, και στις δύο περιπτώσεις φαίνεται τα μοντέλα να αποδίδουν καλύτερα με λιγότερους κανόνες. Αυτό μπορεί να εξηγηθεί, καθώς οι πολλοί κανόνες αυξάνουν τόσο την πολυπλοκότητα, όσο και την πιθανότητα να έχουμε υπερεκπαίδευση. Χειρότερα όλων φαίνεται να αποδίδει το class independent μοντέλο με ακτίνα 0.1, καθώς έχει τις χειρότερες μετρικές. Αξιοσημείωτο είναι το γεγονός πως στη συγκεκριμένη εκτέλεση από την οποία πάρθηκαν όλες οι εν λόγω τιμές των δεικτών απόδοσης, το class independent μοντέλο με ακτίνα 0.9 έχει μηδενικές producer's και user's accuracies για την κλάση 2, αλλά έχει το δεύτερο καλύτερο overall accuracy, και τα δεύτερα καλύτερα producer's και user's accuracies για την κλάση 1. Όπως φαίνεται και από τον πίνακα σφάλματος του εν λόγω μοντέλου, αυτό τείνει να ταξινομεί τα δείγματα αποκλειστικά στην κλάση 1 (ταξινόμησε το 93.62% των χαρακτηριστικών που ανήκουν στην κλάση 1 σωστά, ενώ το 75.86% των χαρακτηριστικών που ταξινόμησε στην κλάση 1, ανήκουν όντως σε αυτή). Το καλό overall accuracy που πετυχαίνει το εν λόγω μοντέλο συμβαίνει πιθανότατα λόγω του ότι η πλειοψηφία των δεδομένων τα οποία ταξινομεί, ανήκουν στην κλάση 1, ωστόσο όλα τα δεδομένα τα οποία ανήκουν στην κλάση 2 ταξινομούνται και αυτά στην 1.

Σχετικά με την επικάλυψη των προβολών των ασαφών συνόλων κάθε cluster στις αντίστοιχες εισόδους όσον αφορά την ενεργοποίηση των κανόνων και γενικότερα την απόδοση του ταξινομητή, παρατηρείται πως τα μοντέλα που εμφανίζουν επικάλυψη, αφενός είναι αυτά με τις μικρότερες τιμές της ακτίνας, συνεπώς αυτά με το μεγαλύτερο αριθμό κανόνων και συναρτήσεων συμμετοχής, αφετέρου τείνουν να έχουν χειρότερη απόδοση (πολυπλοκότητα, υπερεκπαίδευση). Συνεπώς, μια μέθοδος για τη βελτίωση του εν λόγω ζητήματος, είναι η χρήση ενός πιο περιορισμένου αριθμού κανόνων και συναρτήσεων συμμετοχής.

Εφαρμογή σε dataset με υψηλή διαστασιμότητα

Το δεύτερο μέρος της εργασίας υλοποιείται σε κώδικα MATLAB και αποθηκεύεται στο αρχείο main2.m. Σε πρώτη φάση, αφού φορτώνεται και κανονικοποιείται το σετ των δεδομένων που θα χρησιμοποιήσουμε, πρέπει να βρούμε βελτιστοποιημένες τιμές για τις δύο ελεύθερες παραμέτρους που περιλαμβάνει αναγκαστικά το πρόβλημα: τον αριθμό των χαρακτηριστικών προς επιλογή (features number) που θα χρησιμοποιηθούν στην εκπαίδευση των μοντέλων, και την ακτίνα επιρροής των clusters (r_a), η οποία καθορίζει το πλήθος των κανόνων που θα προκύψουν. Ο αριθμός των χαρακτηριστικών που θα εκπαιδεύουμε μπορεί να είναι από 1 μέχρι 178, ενώ η ακτίνα επιρροής παίρνει τιμές από 0 έως 1. Στο σημείο αυτό, πρέπει να επιλέξουμε τόσο τους διαφόρους συνδυασμούς των αριθμών των χαρακτηριστικών και των τιμών της ακτίνας, όσο και πόσους θέλουμε - συμφέρει να εκπαιδεύουμε και να ελέγξουμε σε κάθε εκτέλεση. Οι αριθμοί των χαρακτηριστικών και οι τιμές της ακτίνας για κάθε έναν από αυτούς, αποθηκεύονται σε δύο πίνακες (features_number, r_a _values), και πραγματοποιήθηκαν πολλαπλές εκτελέσεις για διάφορα μεγέθη του κάθε πίνακα, καθώς και για διαφόρους αριθμούς χαρακτηριστικών και τιμές της ακτίνας τη φορά. Το πρώτο συμπέρασμα που προέκυψε, είναι ότι από άποψη χρόνου εκτέλεσης και πολυπλοκότητας, δε συμφέρει η εκπαίδευση για πάνω από τρεις ή τέσσερις διαφορετικούς αριθμούς χαρακτηριστικών σε κάθε εκτέλεση, και ομοίως για τις διαφορετικές τιμές της ακτίνας. Εν συνεχείᾳ, προέκυψε το συμπέρασμα ότι δε συμφέρει ο αριθμός των χαρακτηριστικών να ξεπερνάει τα 15, καθώς ο χρόνος εκτέλεσης αυξάνεται δραστικά, ενώ η απόδοση παρουσιάζει μικρές διαφορές συγκριτικά με μικρότερο αριθμό χαρακτηριστικών (όχι όμως πολύ μικρό). Παρατίθενται ενδεικτικά παρακάτω ορισμένα αποτελέσματα των δοκιμών που γίνανε, και συγκεκριμένα του δείκτη OA, ο οποίος αποτέλεσε και το βασικό κριτήριο για την επιλογή των δύο παραμέτρων που θα χρησιμοποιηθούν στο βέλτιστο, τελικό μοντέλο.

- Δοκιμή 1, για τον αριθμό των χαρακτηριστικών ίσο με 12, 13 και 14, και για τις τιμές της ακτίνας ίσες με 0.3, 0.4 και 0.5 (9 συνδυασμοί). Εκτελέστηκε δύο φορές.

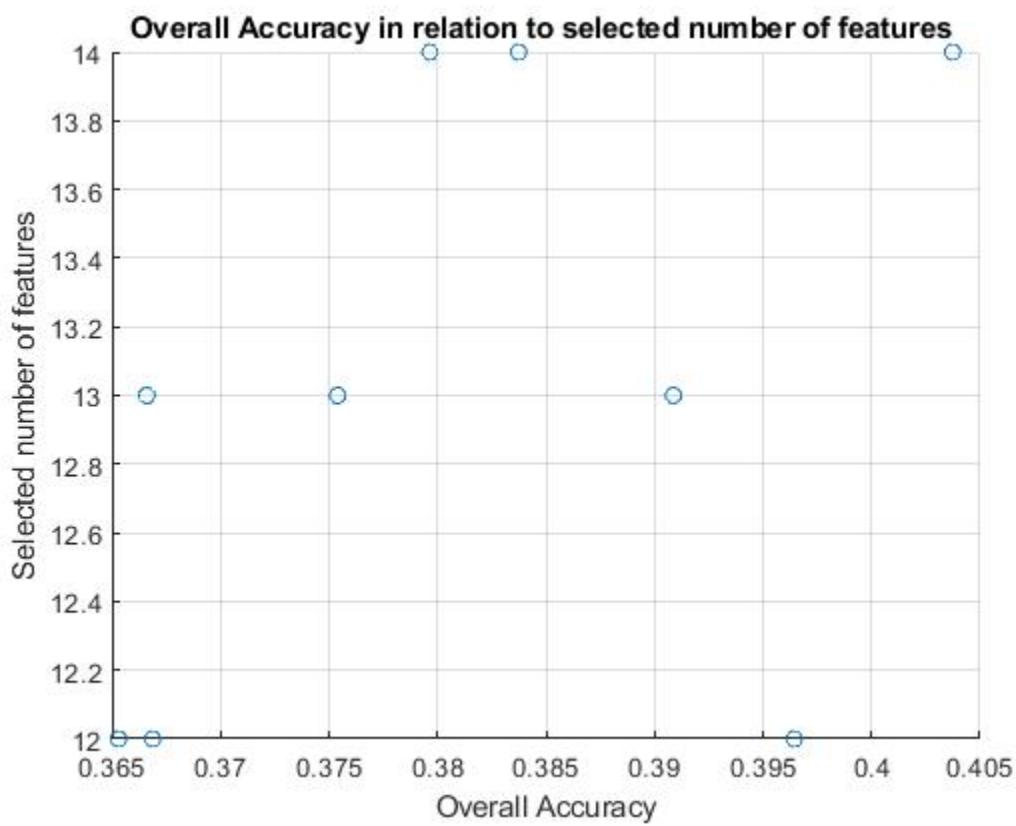
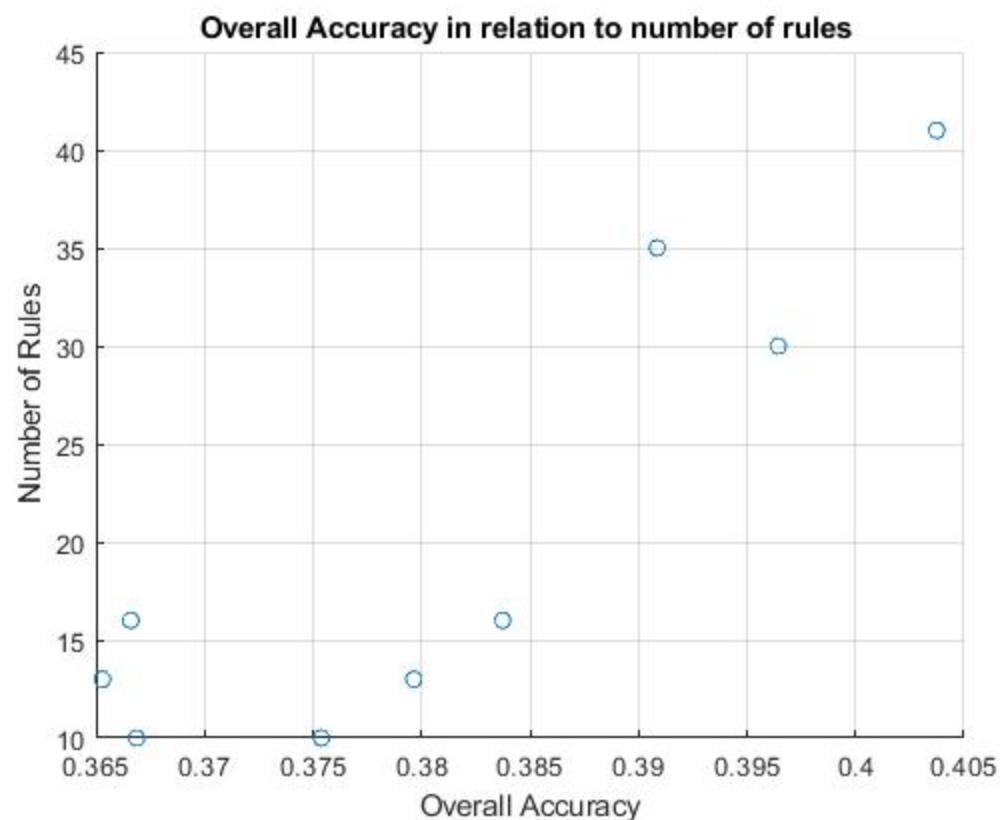
Χρόνος εκτέλεσης ίσος με περίπου 1 ώρα τη φορά.

- Features = 12
 - $r_a = 0.3 \rightarrow OA = 0.391217391304348, 0.396434782608696$, rules = 29, 30
 - $r_a = 0.4 \rightarrow OA = 0.371130434782609, 0.365304347826087$, rules = 14, 13

- $r_a = 0.5 \rightarrow OA = 0.371304347826087, 0.366869565217391$, rules = 11, 10
- Features = 13
 - $r_a = 0.3 \rightarrow OA = 0.388347826086957, 0.390869565217391$, rules = 32, 35
 - $r_a = 0.4 \rightarrow OA = 0.380173913043478, 0.366608695652174$, rules = 14, 16
 - $r_a = 0.5 \rightarrow OA = 0.372434782608696, 0.375391304347826$, rules = 11, 10
- Features = 14
 - $r_a = 0.3 \rightarrow OA = 0.396695652173913, 0.403739130434783$, rules = 36, 41
 - $r_a = 0.4 \rightarrow OA = 0.387304347826087, 0.383739130434783$, rules = 16, 16
 - $r_a = 0.5 \rightarrow OA = 0.377739130434783, 0.379652173913044$, rules = 11, 13

Παρατηρώ πως το overall accuracy σε όλες τις περιπτώσεις είναι καλύτερο για τη μικρότερη ακτίνα από τις τρεις, ενώ λαμβάνει τις υψηλότερες τιμές του για 14 χαρακτηριστικά. Οι επόμενες δύο δοκιμές αποσκοπούν στον έλεγχο του κατά πόσο συμφέρει να μικρίνουμε κι άλλο την τιμή της ακτίνας με παρόμοιο αριθμό χαρακτηριστικών, ή και να μικρίνουμε και αυτό τον αριθμό.

Παρατίθενται επίσης τα διαγράμματα τα οποία να απεικονίζουν την καμπύλη του σφάλματος σε σχέση με τον αριθμό των κανόνων και σε σχέση με τον αριθμό των επιλεχθέντων χαρακτηριστικών.



- Δοκιμή 2, για τον αριθμό των χαρακτηριστικών ίσο με 12, 13 και 14, και για τις τιμές της ακτίνας ίσες με 0.1, 0.2 και 0.3 (9 συνδυασμοί). Δεν μπόρεσε να εκτελεστεί μέσα σε 26 ώρες. Συνεπώς, τόσο μικρή τιμή για την ακτίνα είναι αδύνατο να χρησιμοποιηθεί.
- Δοκιμή 3, για τον αριθμό των χαρακτηριστικών ίσο με 8, 12 και 14, και για τις τιμές της ακτίνας ίσες με 0.2, 0.3 και 0.4 (9 συνδυασμοί). Για τα 8 χαρακτηριστικά, εκτελείται για όλες τις τιμές της ακτίνας σε φυσιολογικά πλαίσια, αλλά δεν παίρνουμε καλύτερο overall accuracy συγκριτικά με τους συνδυασμούς που ελέγχαμε στην πρώτη δοκιμή. Για τα 12 χαρακτηριστικά και ακτίνα ίση με 0.2, ο χρόνος εκτέλεσης αυξάνεται έντονα, ενώ για ακτίνα ίση με 0.3, 0.4 εκτελείται φυσιολογικά όπως είδαμε και στην πρώτη δοκιμή. Για τα 14 ωστόσο χαρακτηριστικά, και ακτίνα ίση με 0.2, ο χρόνος εκτέλεσης σε κάθε fold αυξάνεται υπερβολικά έντονα, και το καθιστά μη εκτελέσιμο.
- Δοκιμή 4, για τον αριθμό των χαρακτηριστικών ίσο με 5, 8 και 12, και για τις τιμές της ακτίνας ίσες με 0.2, 0.3 και 0.4 (9 συνδυασμοί). Επί της ουσίας θέλουμε να δούμε εάν μπορούμε να πάρουμε καλύτερο overall accuracy για ακτίνα ίση με 0.2, αλλά χρησιμοποιώντας αναγκαστικά μικρότερο αριθμό χαρακτηριστικών, λόγω του τι συνέβη στις προηγούμενες δύο δοκιμές.
 - Features = 5
 - r_a = 0.2 -> OA = 0.361478260869565 , rules = 34
 - r_a = 0.3 -> OA = 0.345043478260870, rules = 17
 - r_a = 0.4 -> OA = 0.349217391304348 , rules = 9
 - Features = 8
 - r_a = 0.2 -> OA = 0.383913043478261, rules = 50
 - r_a = 0.3 -> OA = 0.363826086956522 , rules = 18
 - r_a = 0.4 -> OA = 0.367913043478261, rules = 10
 - Features = 12
 - r_a = 0.2 -> OA = 0.397217391304348, rules = 493
 - r_a = 0.3 -> OA = 0.398869565217391, rules = 26
 - r_a = 0.4 -> OA = 0.3740000000000000, rules = 14

Οι 493 κανόνες που εμφανίζονται για 12 features και ακτίνα ίση με 0.2 καθιστούν απαγορευτική την επιλογή τόσο μικρής τιμής για την ακτίνα, ενώ παράλληλα βλέπουμε πως το overall accuracy όχι απλά δεν είναι βελτιώθηκε, αλλά είναι και οριακά χειρότερο από ορισμένα αποτελέσματα της πρώτης δοκιμής. Έχουμε λοιπόν πλέον, αρκετά δεδομένα ώστε να επιλέξουμε τον αριθμό των χαρακτηριστικών και την τιμή της ακτίνας για το τελικό, βέλτιστο μοντέλο που θα εκπαιδεύσουμε.

Όσον αφορά την υλοποίηση των παραπάνω, για κάθε διαφορετικό αριθμό χαρακτηριστικών, και για κάθε διαφορετική τιμή της ακτίνας, διαχωρίζουμε τα δεδομένα σε ποσοστό 80% για εκπαίδευση και 20% για έλεγχο, και ορίζουμε έναν πίνακα για να αποθηκεύσουμε το overall accuracy που θα προκύψει μέσα στο cross validation. Εν συνέχεια προχωρούμε στο cross validation, το οποίο είναι 5-fold, και μέσα στο οποίο διαχωρίζω το 80% των δεδομένων εκπαίδευσης από προηγουμένως, σε ποσοστό 75% για εκπαίδευση, και 25% για επικύρωση. Έτσι, το αρχικό σετ δεδομένων έχει διαχωριστεί σε ποσοστό 60% δεδομένα εκπαίδευσης (trnData), 20% δεδομένα επικύρωσης, και 20% δεδομένα ελέγχου (tstData), όπως ορίζει η εκφώνηση. Επίσης, δημιουργούμε 5 clusters, ένα για κάθε κλάση που έχουμε. Εν συνεχείᾳ εκπαιδεύονται πέντε διαφορετικά μοντέλα με τις ίδιες παραμέτρους (5-fold cross validation), για 100 epochs το καθένα, και υπολογίζεται το overall accuracy του καθενός, το οποίο αποθηκεύεται στη συνέχεια στον πίνακα που ορίσαμε στην αρχή. Τέλος, μετά το πέρας του cross validation, υπολογίζουμε το μέσο όρο του overall accuracy, το οποίο έχει λάβει πέντε διαφορετικές τιμές κατά τη διάρκεια, και τον αποθηκεύουμε στον τελικό πίνακα όπου θα έχουμε όλα τα overall accuracies για κάθε συνδυασμό αριθμού χαρακτηριστικών και τιμών της ακτίνας που επιλέξαμε στην αρχή.

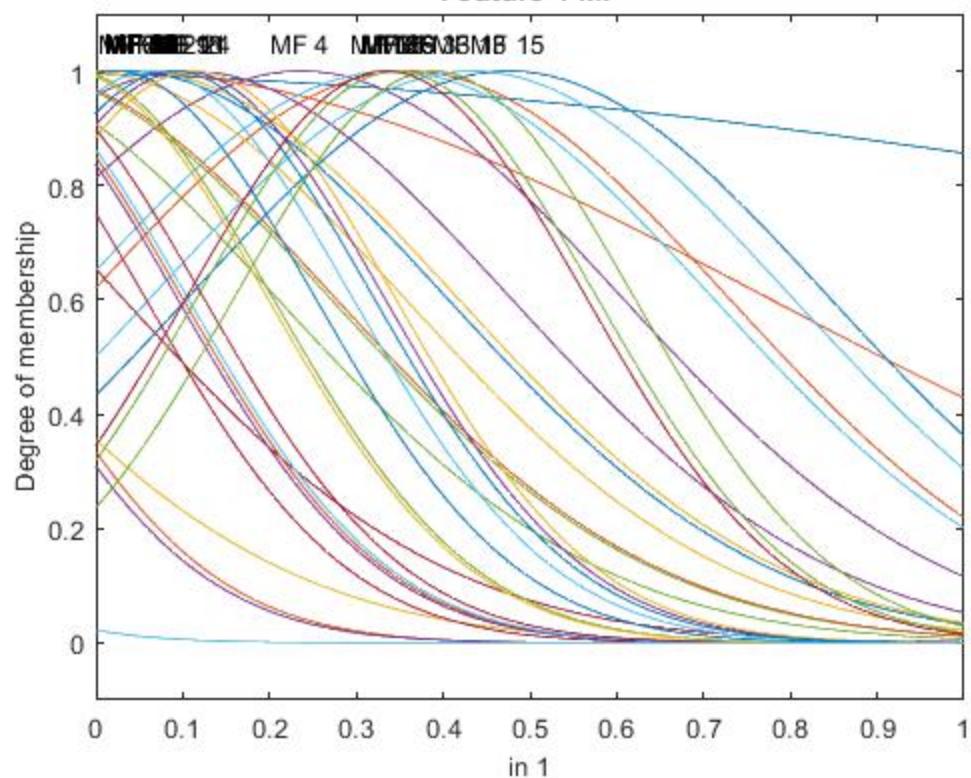
Εκπαίδευση του τελικού, βέλτιστου μοντέλου

Το τελευταίο μέρος της εργασίας υλοποιείται σε κώδικα MATLAB και αποθηκεύεται στο αρχείο main3.m, όπου και εκπαιδεύομε το τελικό, βέλτιστο μοντέλο χρησιμοποιώντας τον αριθμό των χαρακτηριστικών και την τιμή της ακτίνας που επιλέχθηκαν βάση των συμπερασμάτων του δευτέρου μέρους της εργασίας. Ο αριθμός των χαρακτηριστικών που επιλέχθηκε είναι 14, ενώ η τιμή της ακτίνας είναι 0.3. Το μοντέλο εκπαιδεύεται ακριβώς με τον ίδιο τρόπο όπως τα προηγούμενα, στη συνέχεια υπολογίζουμε και αποθηκεύουμε το μέσο όρο των δώδεκα δεικτών απόδοσης που μας ενδιαφέρουν και προκύπτουν από τη διαδικασία του cross validation, και δημιουργούμε ενδεικτικά μερικά

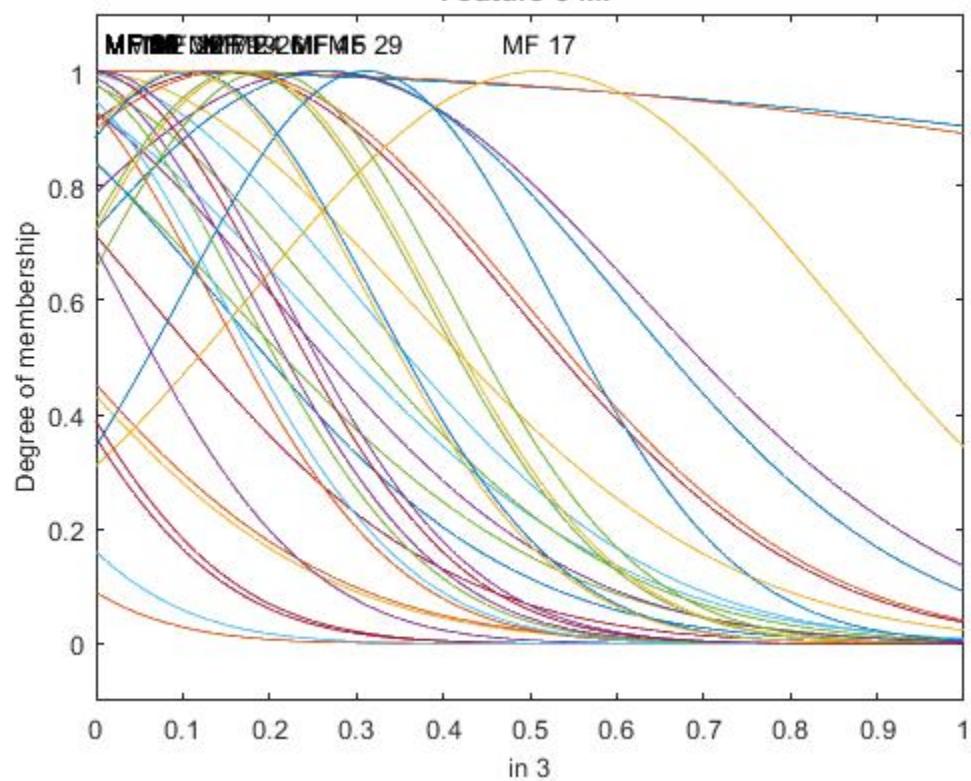
διαγράμματα συναρτήσεων συμμετοχής στην αρχική και στην τελική τους μορφή, τα διαγράμματα όπου αποτυπώνονται οι προβλέψεις του τελικού μοντέλου και οι πραγματικές τιμές, το διάγραμμα εκμάθησης, όπου απεικονίζεται το σφάλμα συναρτήσει του αριθμού των επαναλήψεων, και τα οποία παρατίθενται παρακάτω, μαζί με τον πίνακα των τιμών των δώδεκα δεικτών απόδοσης.

Συναρτήσεις συμμετοχής για διαφορετικά χαρακτηριστικά πριν την εκπαίδευση

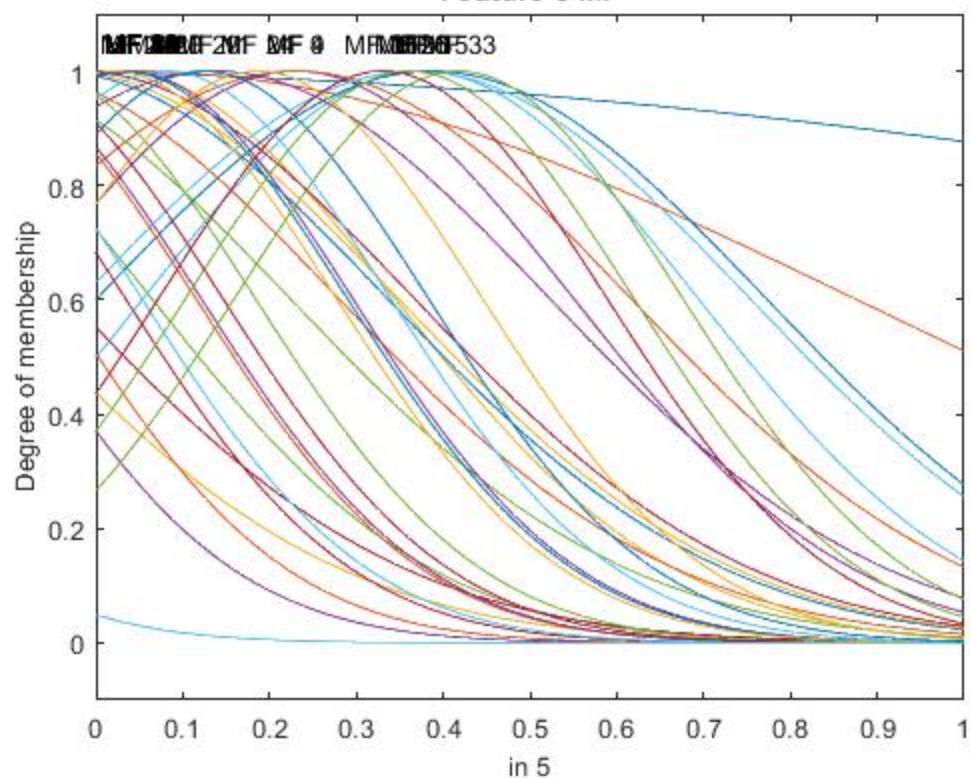
Feature 1 MF



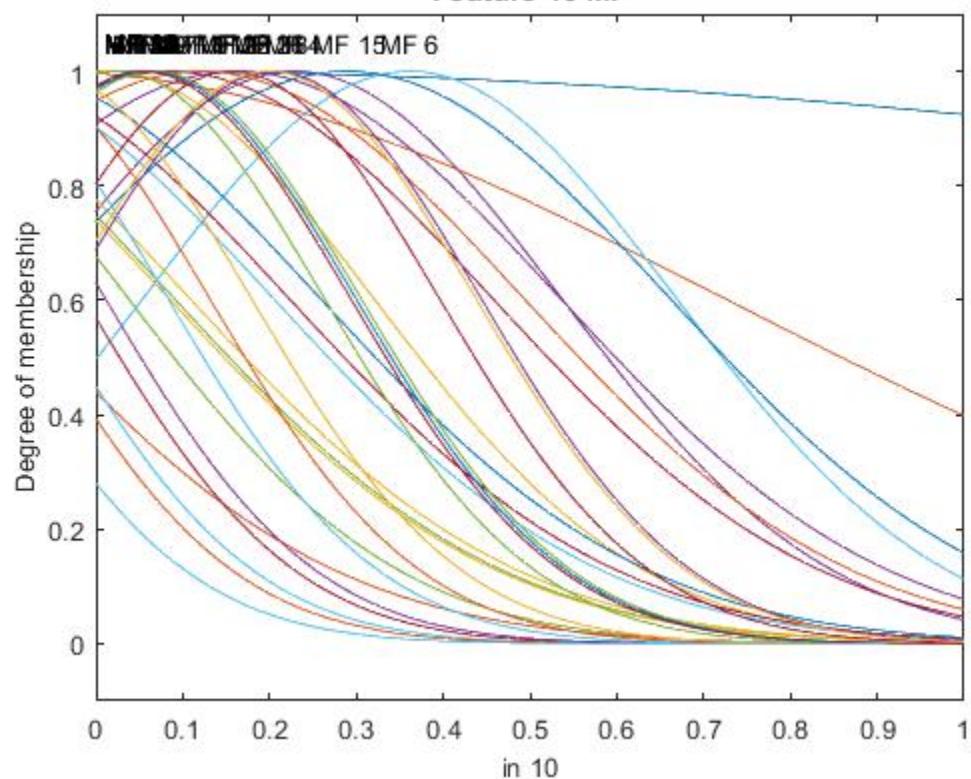
Feature 3 MF



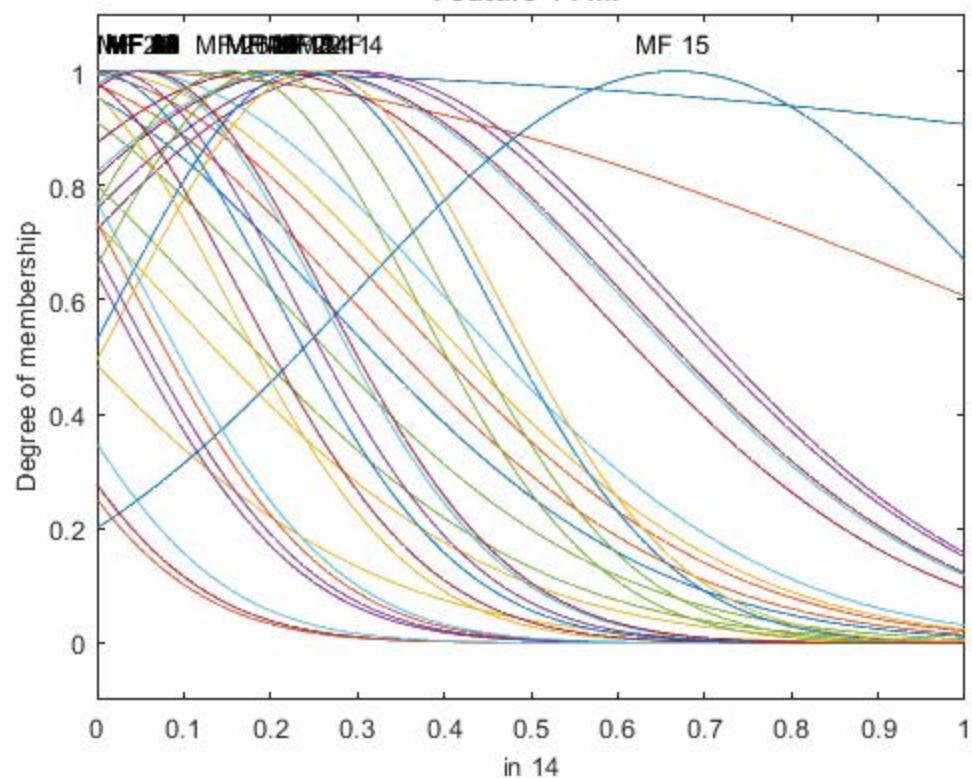
Feature 5 MF



Feature 10 MF

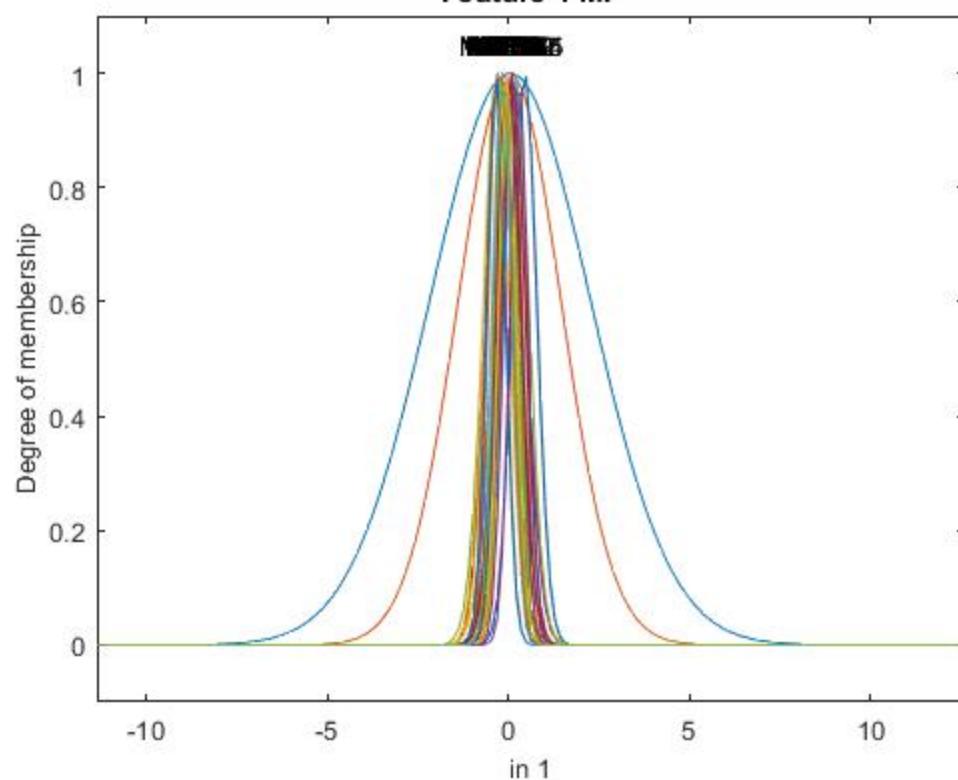


Feature 14 MF

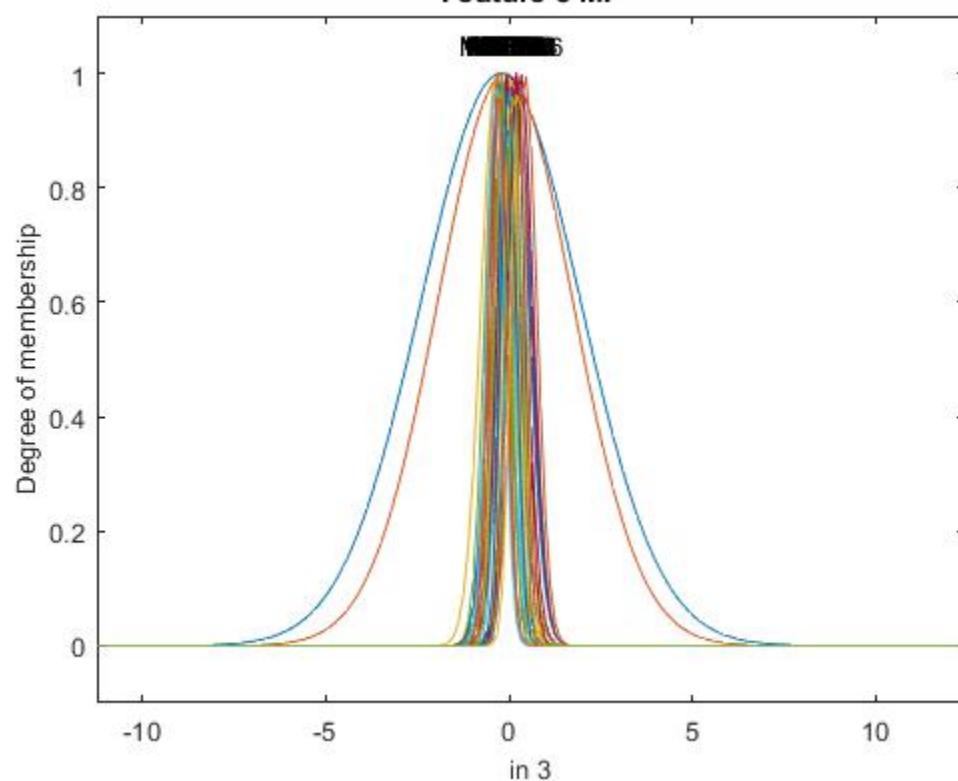


Συναρτήσεις συμμετοχής για τα ίδια χαρακτηριστικά μετά την εκπαίδευση

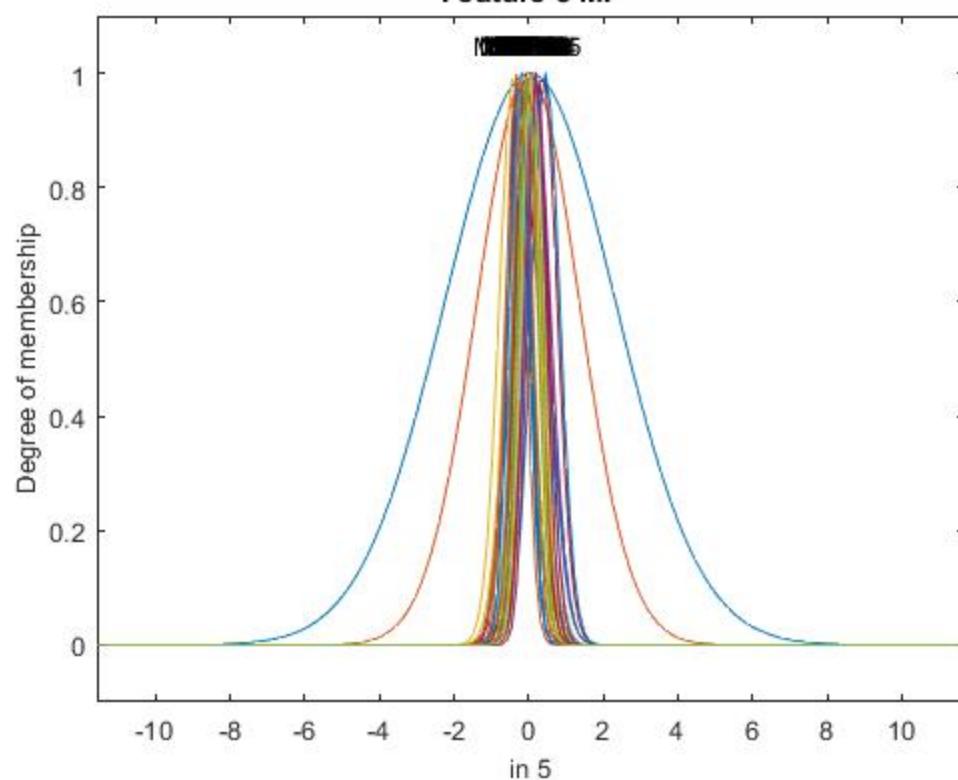
Feature 1 MF



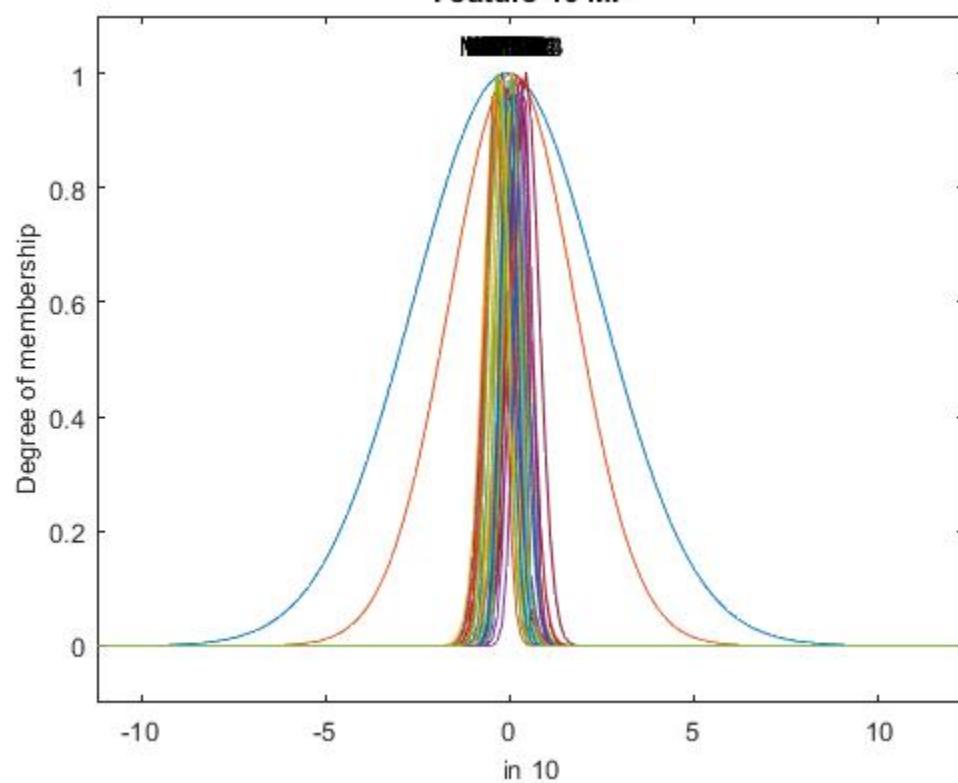
Feature 3 MF

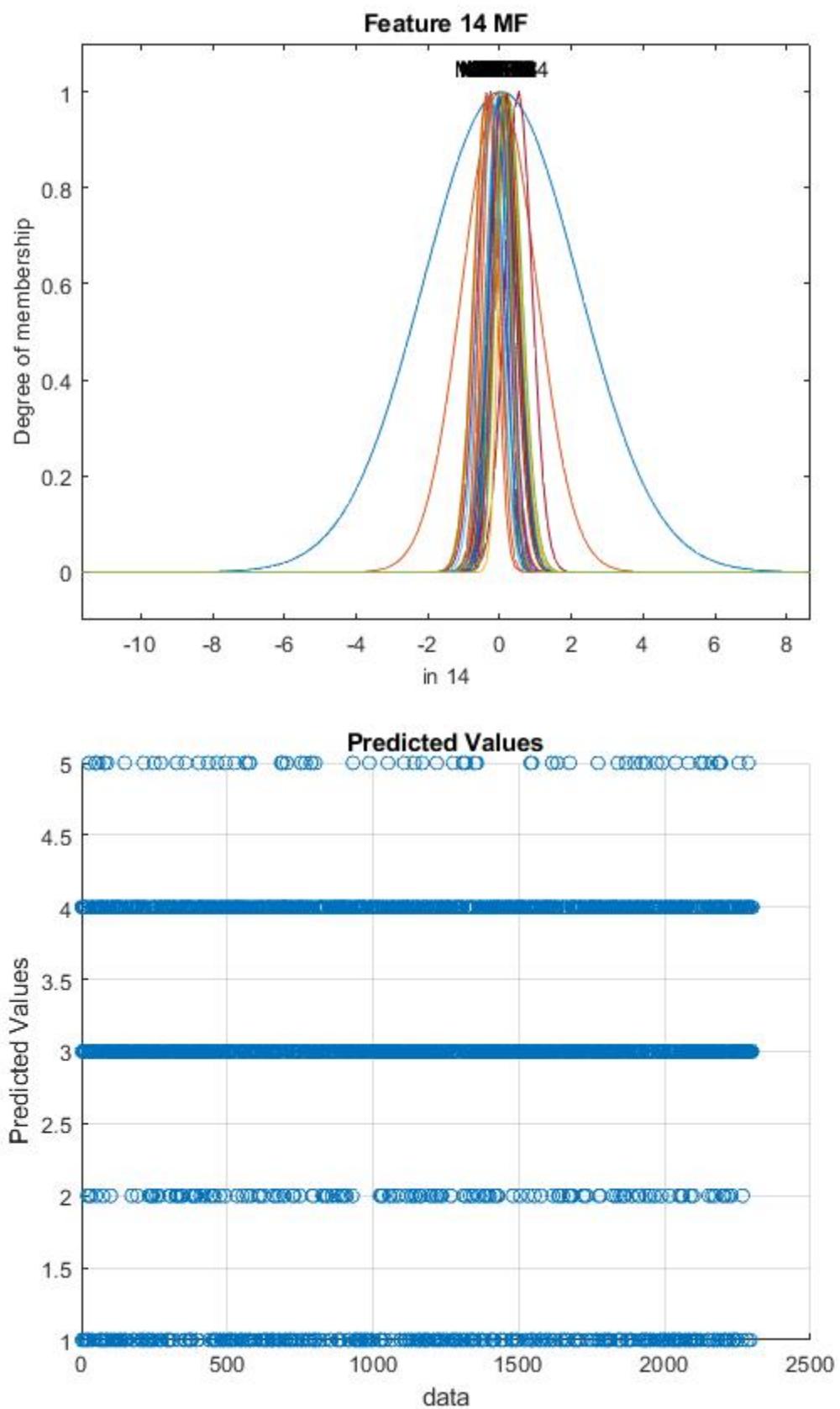


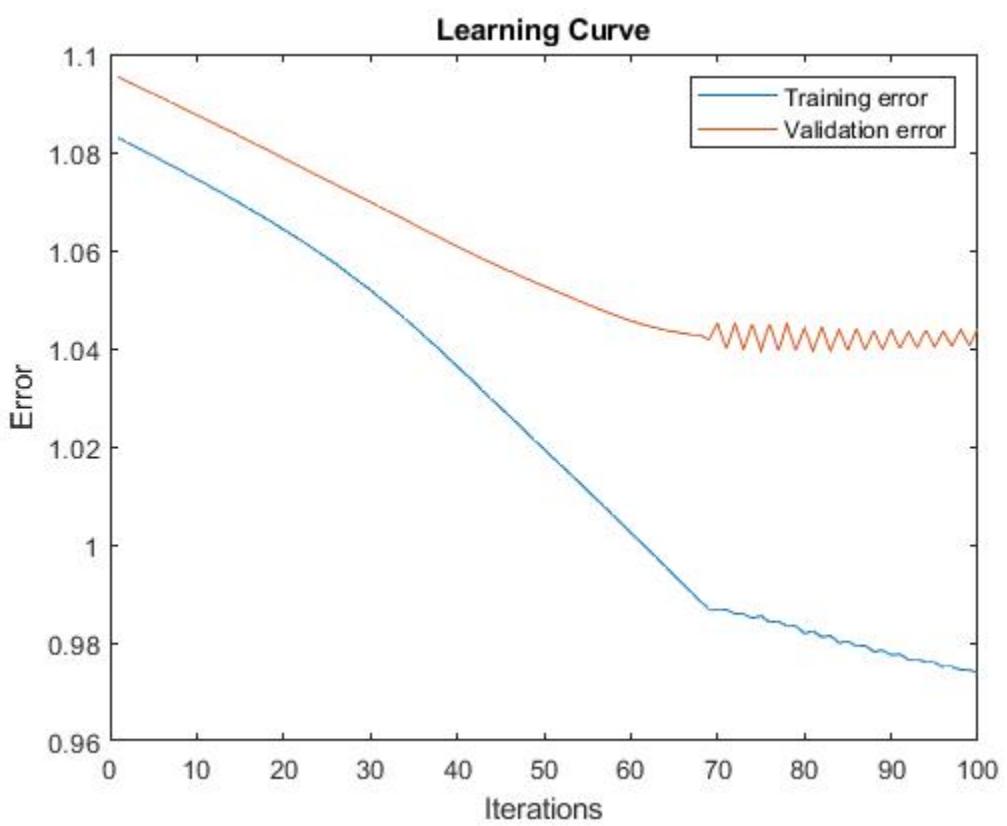
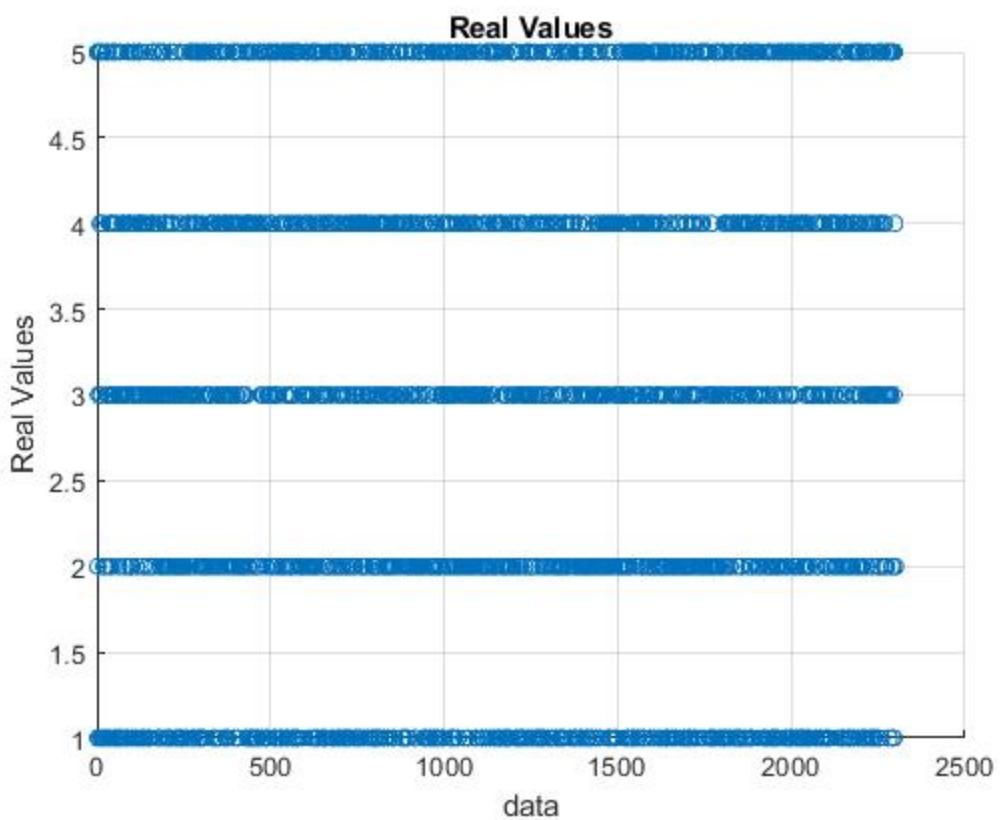
Feature 5 MF



Feature 10 MF







Πίνακας σφαλμάτων ταξινόμησης (από το τελευταίο μοντέλο του cross validation)

328	69	54	9	0
25	44	271	119	1
1	21	303	129	6
2	33	186	213	25
0	4	152	266	38

Τιμές δεικτών απόδοσης (οι μέσοι όροι των τιμών των 5 μοντέλων του cross validation)

OA	0.4011
K	0.2514
PA for class 1	0.7496
PA for class 2	0.0704
PA for class 3	0.6457
PA for class 4	0.4643
PA for class 5	0.0757
UA for class 1	0.9330
UA for class 2	0.2078
UA for class 3	0.3050
UA for class 4	0.2898
UA for class 5	0.5

Βλέποντας το τον πίνακα σφαλμάτων ταξινόμησης και τους δείκτες απόδοσης του τελικού μοντέλου, μπορούμε να εξάγουμε τα εξής συμπεράσματα:

- Το μοντέλο ταξινόμησε σωστά περίπου το 75% των δεδομένων που ανήκουν στην κλάση 1, ενώ το 93% των δεδομένων τα οποία ταξινόμησε στην εν λόγω κλάση, ανήκουν όντως σε αυτή.
- Ταξινόμησε σωστά μόνο το 7% των δεδομένων που ανήκουν στην κλάση 2, ενώ μόνο το 20% των δεδομένων τα οποία ταξινόμησε στην εν λόγω κλάση, ανήκουν όντως σε αυτή. Τα περισσότερα δεδομένα που ανήκουν στην κλάση 2 ταξινομήθηκαν στις 3 και 4.
- Ταξινόμησε σωστά περίπου το 65% των δεδομένων που ανήκουν στην κλάση 3, αλλά μόνο το 30% των δεδομένων τα οποία ταξινόμησε στην εν λόγω κλάση, ανήκουν όντως σε αυτή.
- Ταξινόμησε σωστά περίπου το 46% των δεδομένων που ανήκουν στην κλάση 4, και μόνο σχεδόν το 30% των δεδομένων τα οποία ταξινόμησε στην εν λόγω κλάση, ανήκουν όντως σε αυτή.
- Ταξινόμησε σωστά μόνο το σχεδόν 8% των δεδομένων που ανήκουν στην κλάση 5, ενώ το 50% των δεδομένων τα οποία ταξινόμησε στην εν λόγω κλάση, ανήκουν όντως σε αυτή. Τα περισσότερα δεδομένα που ανήκουν στην κλάση 5 ταξινομήθηκαν στις 3 και 4.

Συνεπώς:

- Το μοντέλο ταξινομεί πολύ καλά τα δεδομένα τα οποία ανήκουν πράγματι στην κλάση 1, και ταυτόχρονα ταξινομεί ελάχιστα δεδομένα άλλων κλάσεων στην κλάση 1.
- Ταξινομεί από μέτρια έως σχετικά καλά τα δεδομένα τα οποία ανήκουν πράγματι στις κλάσεις 3, και 4, αλλά ταξινομεί στις εν λόγω κλάσεις πολλά δεδομένα τα οποία δεν ανήκουν σε αυτές.
- Δεν ταξινομεί σωστά την πλειοψηφία των δεδομένων που ανήκουν στις κλάσεις 2 και 5, ενώ ταξινομεί ταυτόχρονα σε αυτές τις κλάσεις σημαντικό έως μεγάλο ποσοστό δεδομένων που δεν ανήκουν σε αυτές.
- Τα περισσότερα δεδομένα που ανήκουν στις κλάσεις 2 και 5 ταξινομούνται στις κλάσεις 3 και 4.
- Εκτός των δεδομένων που ανήκουν στην κλάση 1, η πλειοψηφία των υπολοίπων δεδομένων ταξινομούνται στις κλάσεις 3 και 4.
- Τα δύο παραπάνω υποθέτουμε ότι συμβαίνουν επειδή οι προβλέψεις τείνουν προς το μέσο όρο των κλάσεων, ώστε να ελαχιστοποιείται το σφάλμα.

Καταφέραμε να επιτύχουμε overall accuracy κοντά στο 0.4, που συγκριτικά με όλες τις δοκιμές για τους διαφόρους αριθμούς χαρακτηριστικών και τιμών της ακτίνας που πραγματοποιήθηκαν προηγουμένως, είναι μία από τις υψηλότερες τιμές. Επίσης, παρατηρούμε μια αρκετά υψηλή τιμή στο K, πάντα συγκριτικά με τις προηγούμενες δοκιμές και εκτελέσεις. Ο αριθμός των κανόνων του ασαφούς συστήματος συμπερασμού είναι στην προκειμένη ίσος με 39, πολύ κοντά σε αυτό που θα περιμέναμε, καθώς είναι πολύ κοντά στους αριθμούς των κανόνων των διαφόρων μοντέλων που δοκιμάστηκαν, και τα οποία μας έδωσαν ικανοποιητικό overall accuracy. Αν για το ίδιο πλήθος χαρακτηριστικών, είχαμε επιλέξει grid partitioning με δύο ή τρία ασαφή σύνολα ανά είσοδο, τότε ο αριθμός των κανόνων θα εκτοξευόταν σε 2^{14} ή 3^{14} . Το γεγονός αυτό καθιστά όχι μονο απαγορευτική την εν λόγω εφαρμογή και την εκτέλεσή της πρακτικά, αλλά θα είχε και πολύ μεγάλες πιθανότητες να εμφανίσει υπερεκπαίδευση. Συνεπώς, ο διαμερισμός του συνολικού χώρου εισόδου έχει πολύ σημαντική επίδραση στο ποσοστό των ενεργών κανόνων, και κατά συνέπεια στη δυνατότητα εκτέλεσης - εκπαίδευσης, αλλά και στην αξιοπιστία του μοντέλου.