Εργασία
Στα Ασαφή Συστήματα
(Group4_Ser6)

ΟΝΟΜΑ: Αντώνης
ΕΠΩΝΥΜΟ: Μυρσινιάς
ΑΕΜ:8873
ΕΤΟΣ:2020

Περιεχόμενα

Πρόβλημα

Στόχος της εργασίας αυτής είναι να διερευνηθεί η ικανότητα των μοντέλων TSK στην επίλυση προβλημάτων ταξινόμησης (classification). Συγκεκριμένα, επιλέγονται δύο σύνολα δεδομένων από το UCI repository με σκοπό την ταξινόμηση, από τα διαθέσιμα δεδομένα δειγμάτων, στις εκάστοτε κλάσεις τους, με χρήση ασαφών νευρικών μοντέλων. Η εργασία αποτελείται από δύο μερη, το πρώτο από τα οποία προορίζεται για μια απλή διεύρυνση της διαδικασίας εκπαίδευσης και αξιολόγησης των TSK μοντέλων, ενώ το δεύτερο περιλαμβάνει μια πιο συστηματική προσέγγιση στο πρόβλημα της εκμάθησης από δεδομένα, σε συνδυασμό με προεπεξεργαστικά βήματα όπως επιλογή χαρακτηριστικών (feature selection) και μεθόδους βελτιστοποίησης των μοντέλων μέσω της διασταυρωμένης επικύρωσης (cross validation).

Εφαρμογή σε απλό dataset

Στην πρώτη φάση της εργασίας, επιλέγεται από το UCI repository το avila dataset, το οποίο περιλαμβάνει 20876 δείγματα (instances), από 10 χαρακτηριστικά (attributes) το καθένα. Ακολουθούνται τα παρακάτω βήματα.

Διαχωρισμός σε σύνολα εκπαίδευσης-επικύρωσης-ελέγχου

Αρχικά είναι απαραίτητος ο διαχωρισμός του συνόλου δεδομένων σε τρία μη επικαλυπτόμενα υποσύνολα $\{D_{trn}, D_{val}, D_{chk}\}$ ως εξής:

- D_{trn} = 60% του αρχικού σετ (Σετ το οποίο θα χρησιμοποιηθεί κατά την εκπαίδευση του μοντέλου).
- D_{val} = 20% του αρχικού σετ (Σετ το οποίο θα χρησιμοποιηθεί για την επικύρωση και αποφυγή του φαινομένου της υπερεκπαίδευσης).
- D_{chk} = 20% του αρχικού σετ (Σετ που θα χρησιμοποιηθεί για τον έλεγχο της απόδοσης του τελικού μοντέλου).

Για να επιτύχουμε καλή απόδοση, θα πρέπει η συχνότητα εμφάνισης δειγμάτων που ανήκουν σε μία συγκεκριμένη κλάση, σε κάθε ένα από τα τρία σύνολα διαμέρισης, να είναι όσο το δνατόν πιο «όμοια» με την ανίστοιχη συχνότητα εμφάνισης τους στο αρχικό σύνολο δεδομένων. Πιο συγκεκριμένα, στους παρακάτω πίνακες φαίνεται η συχνότητα της κάθε κλάσης στο κάθε σύνολο.

Συχνότητα κάθε κλάσεις στο Συνολικό Dataset:

Class	Count	Percent
1	8572	41.08%
2	10	0.05%
3	206	0.99%
4	705	3.38%
5	2190	10.50%
6	3923	18.80%
7	893	4.28%
8	1039	4.98%
9	1663	7.97%
10	89	0.43%
11	1044	5.00%
12	533	2.55%

Πίνακας 1: συχνότητα κλάσεων στο συνολικό Dataset

Συχνότητα κάθε κλάσεις στο σετ εκπαίδευσης:

Class	Count	Percent
1	5143	41.08%
2	6	0.05%
3	124	0.99%
4	423	3.38%
5	1314	10.50%
6	2354	18.80%
7	536	4.28%
8	623	4.98%

9	998	7.97%
10	53	0.42%
11	626	5.00%
12	320	2.56%

Πίνακας 2: συχνότητα κλάσεων στο σετ εκπαίδευσης

Συχνότητα κάθε κλάσεις στο σετ επικύρωσης:

Class	Count	Percent
1	1714	41.05%
2	2	0.05%
3	41	0.98%
4	141	3.38%
5	438	10.49%
6	785	18.80%
7	179	4.29%
8	208	4.98%
9	333	7.98%
10	18	0.43%
11	209	5.01%
12	107	2.56%

Πίνακας 3: συχνότητα κλάσεων στο σετ επικύρωσης

Συχνότητα κάθε κλάσεις στο σετ ελέγχου:

Class	Count	Percent
1	1715	41.11%
2	2	0.05%
3	41	0.98%
4	141	3.38%
5	438	10.50%
6	784	18.79%
7	178	4.27%

8	208	4.99%
9	332	7.96%
10	18	0.43%
11	209	5.01%
12	106	2.54%

Πίνακας 4: συχνότητα κλάσεων στο σετ εκπαίδευσης

Εκπαίδευση TSK μοντέλων με διαφορετικές παραμέτρους

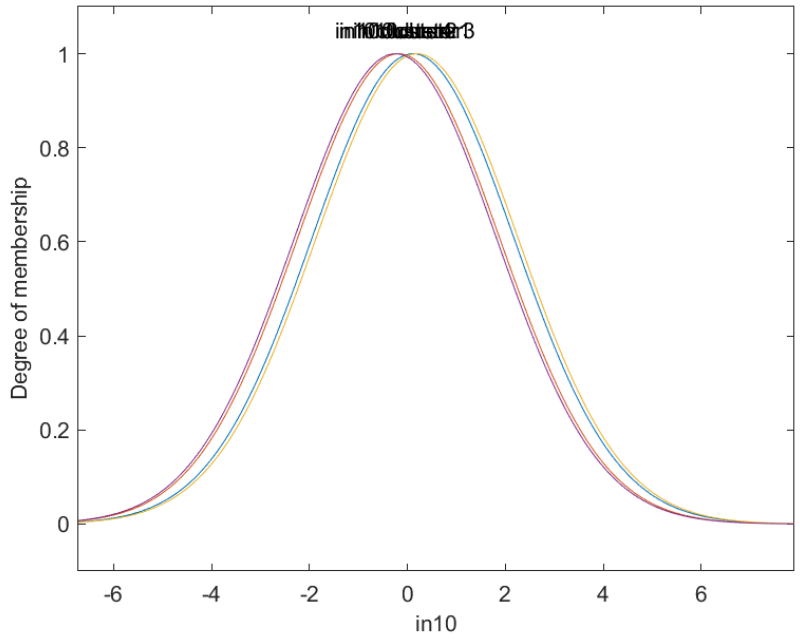
Τα μοντέλα εκπαιδεύτηκαν με την υβριδική μέθοδο, σύμφωνα με την οποία οι παράμετροι των συναρτήσεων συμμετοχής βελτιστοποιούνται μέσω της μεθόδου της οπισθοδιάδοσης (backpropagation algorithm). Η διαμέριση του χώρου εισόδου έγινε με την μέθοδο Subtractive Clustering (SC) και τα TSK μοντέλα που προέκυψαν διαφέρουν ως προς την παράμετρο που καθορίζει τον αριθμό των κανόνων. Επιπλέον, εφόσον η έξοδος αποτελείται από έναν ακέραιο αριθμό, ενδεικτικό της κλάσης στην οποία ανήκει το εκάστοτε δείγμα, έγινε χειροκίνητη του τύπου της συνάρτησης εξόδου από linear σε constant. Συνολικά, έχουμε πέντε διαφορευικά μοντέλα, των οποίων τα χαρακτιρηστικά παρουσιάζονται στον παρακάτω πίνακα.

Model	Range	Squash	Rules
1	0.4	0.85	4
2	0.24	0.7	8
3	0.22	0.5	12
4	0.4	0.45	16
5	0.58	0.42	20

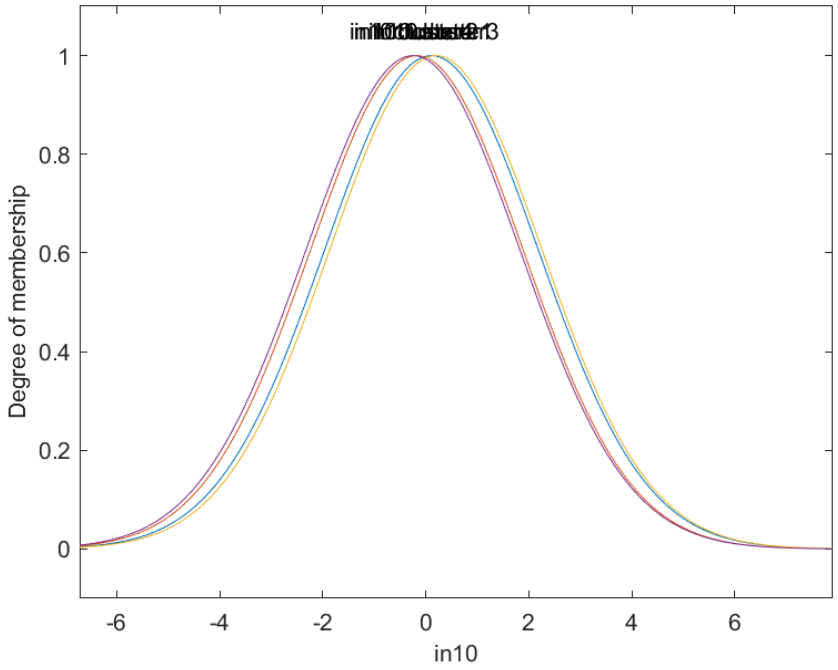
Πίνακας 5: Χρακτηριστικά μοντέλων

Μοντέλο 1

Στα παρακάτω σχήματα παρουσιάζονται οι συναρτήσεις συμμετοχής πριν και μετά την εκπαίδευση.

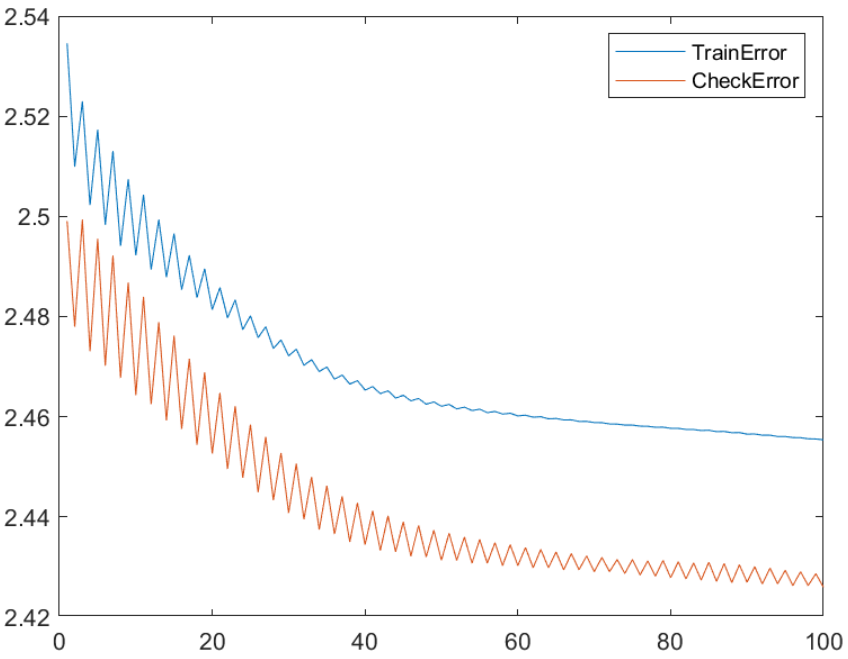


Σχήμα 1: Συναρτήσεις συμμετοχής πριν την εκπαίδευση (Μοντέλο 1)

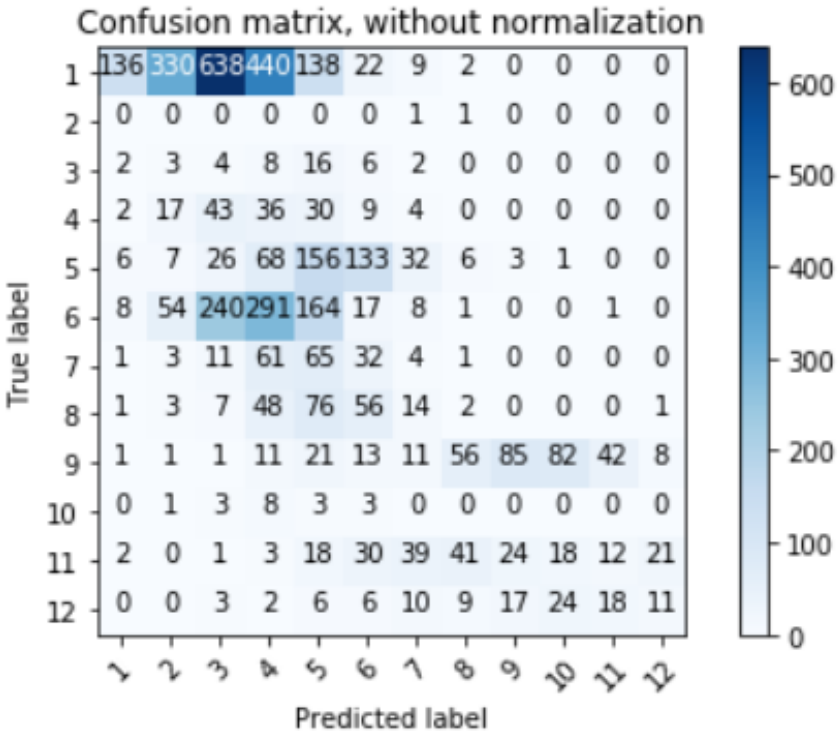


Σχήμα 2: Συναρτήσεις συμμετοχής μετά την εκπαίδευση (Μοντέλο 1)

Στη συνέχεια, παρουσιάζονται οι καμπύλες εκμάθησης (learning curves) και ο error matrix.



Σχήμα 3: Καμπύλες εκμάθησης(Μοντέλο 1)



Σχήμα 4: Error Matrix (Μοντέλο 1)

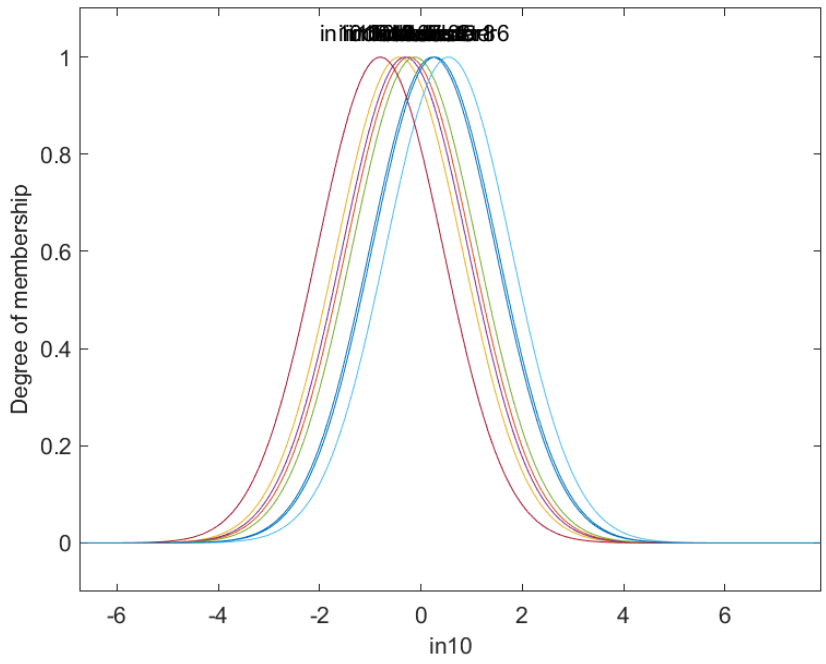
Στον παρακάτο πίνακα παρουσιάζονται οι μετρικές OA, PA, UA, $\hat{k}$.

Class	PA	UA	OA	$\hat{k}$
1	0.8553	0.1504	0.111	0.0496
2	0	0.5000		
3	0.0041	0.1463		
4	0.0369	0.1986		
5	0.2251	0.2991		
6	0.0520	0.0383		
7	0.0299	0.0112		
8	0.0168	0.0096		
9	0.6589	0.4307		
10	0	0		
11	0.1644	0.1388		
12	0.2683	0.1792		

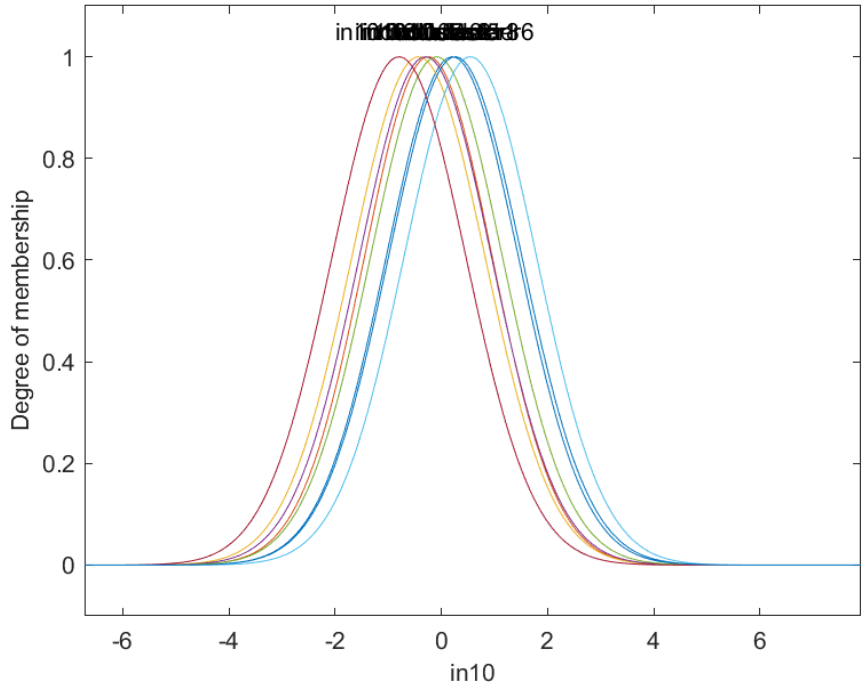
Πίνακας 6: Μετρικές OA, PA, UA, $\hat{k}$

Μοντέλο 2

Στα παρακάτω σχήματα παρουσιάζονται οι συναρτήσεις συμμετοχής πριν και μετά την εκπαίδευση.

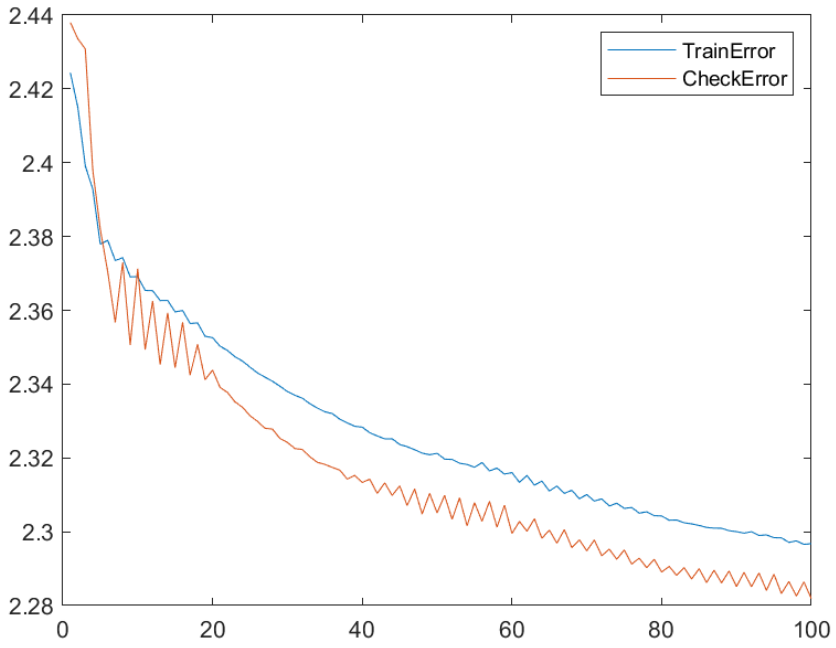


Σχήμα 5: Συναρτήσεις συμμετοχής πριν την εκπαίδευση (Μοντέλο 2)

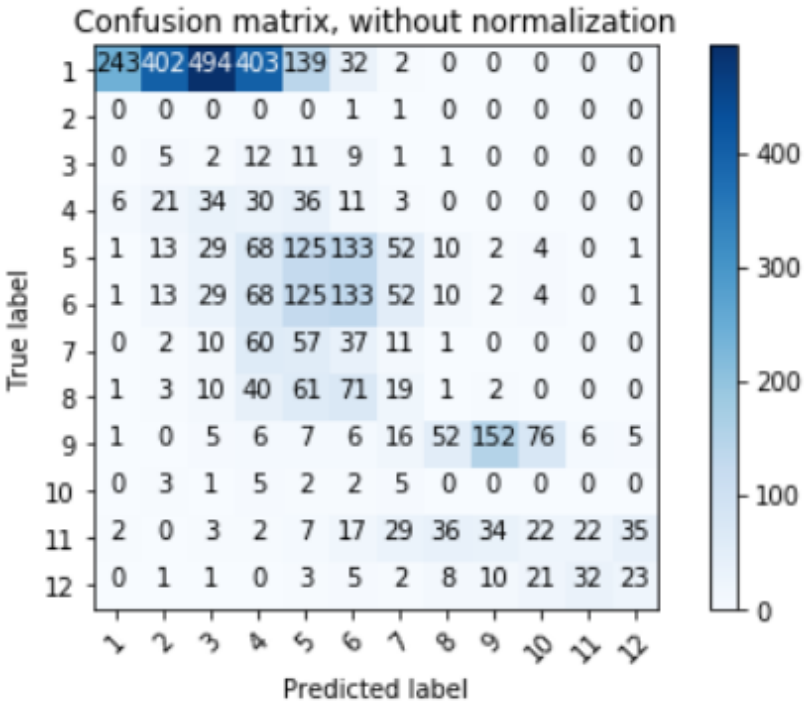


Σχήμα 6: Συναρτήσεις συμμετοχής μετά την εκπαίδευση (Μοντέλο 2)

Στη συνέχεια, παρουσιάζονται οι καμπύλες εκμάθησης (learning curves) και ο error matrix.



Σχήμα 7: Καμπύλες εκμάθησης(Μοντέλο 2)



Σχήμα 8: Error Matrix (Μοντέλο 2)

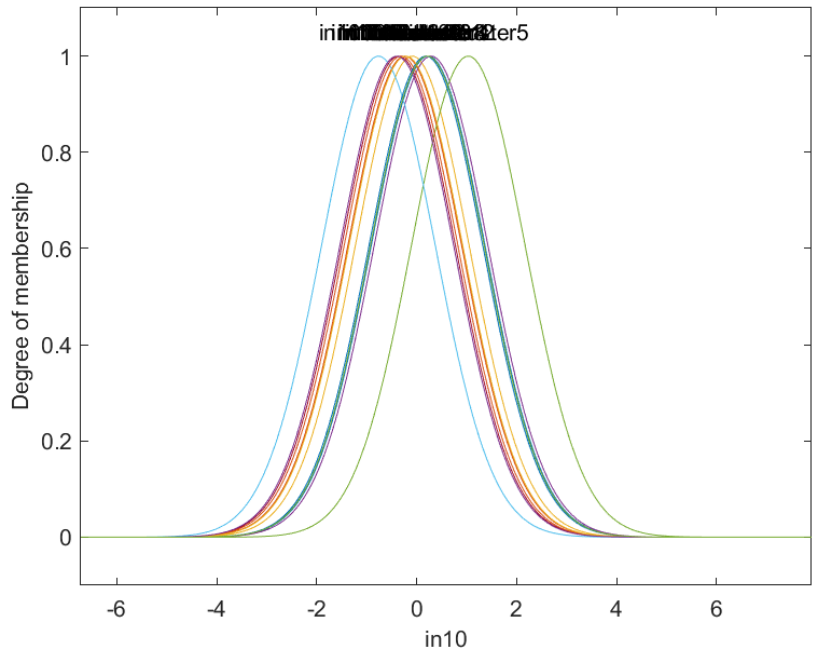
Στον παρακάτο πίνακα παρουσιάζονται οι μετρικές OA, PA, UA, $\hat{k}$.

Class	PA	UA	OA	$\hat{k}$
1	0.8836	0.1417	0.1558	0.0863
2	0	0		
3	0.0025	0.0488		
4	0.0340	0.2128		
5	0.2019	0.2854		
6	0.1123	0.0523		
7	0.0759	0.0618		
8	0.0091	0.0048		
9	0.7600	0.4578		
10	0	0		
11	0.3667	0.1053		
12	0.3594	0.2170		

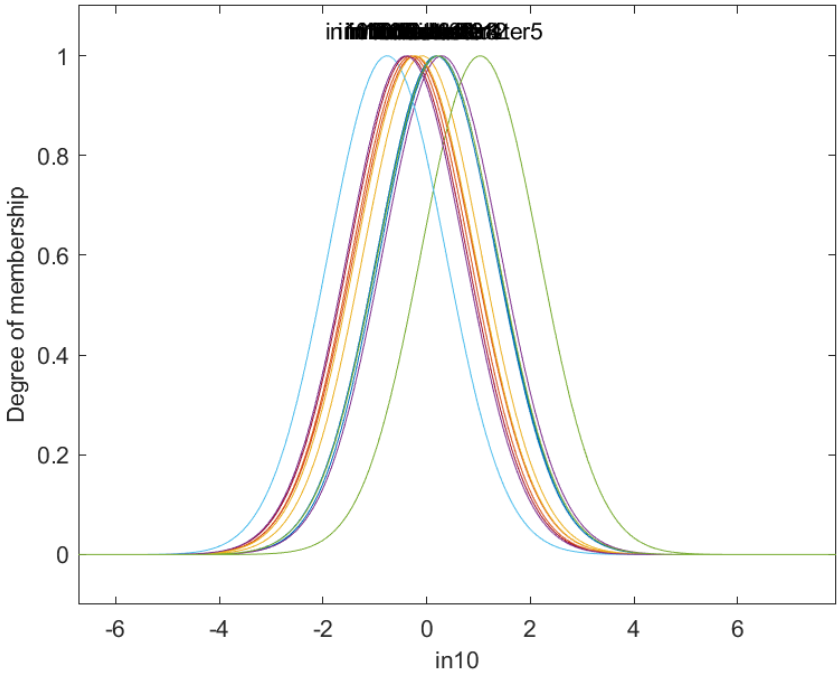
Πίνακας 6: Μετρικές OA, PA, UA, $\hat{k}$

Μοντέλο 3

Στα παρακάτω σχήματα παρουσιάζονται οι συναρτήσεις συμμετοχής πριν και μετά την εκπαίδευση.

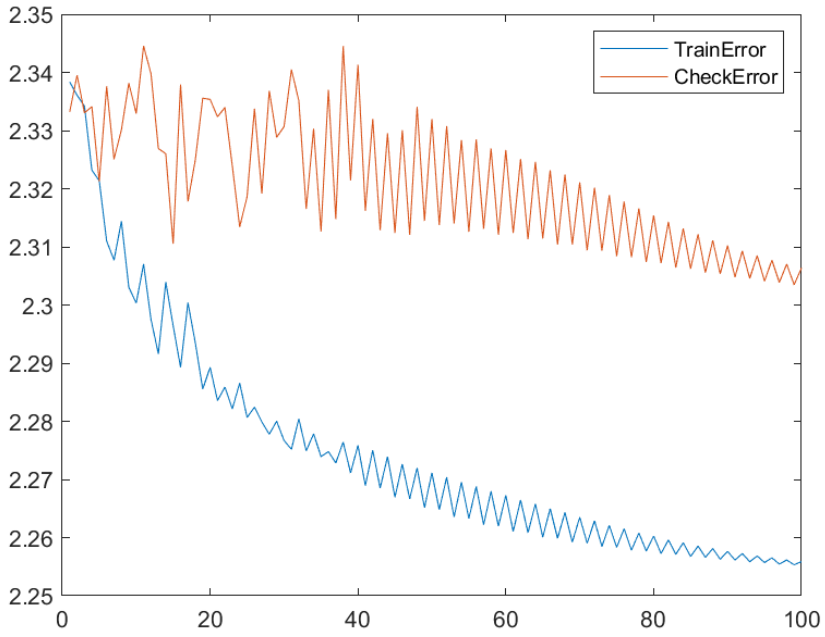


Σχήμα 9: Συναρτήσεις συμμετοχής πριν την εκπαίδευση (Μοντέλο 3)

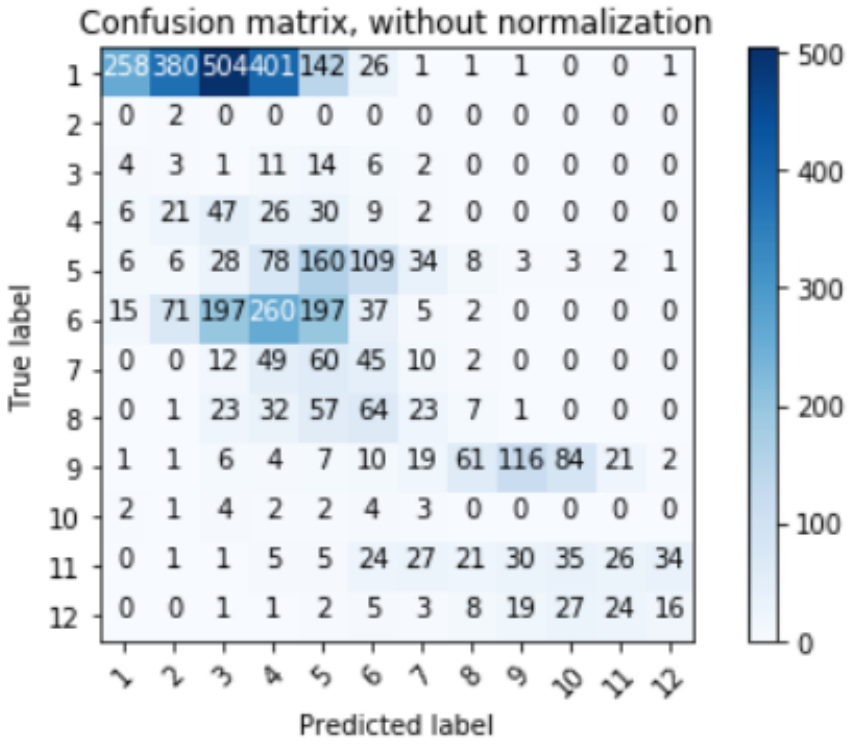


Σχήμα 10: Συναρτήσεις συμμετοχής μετά την εκπαίδευση (Μοντέλο 3)

Στη συνέχεια, παρουσιάζονται οι καμπύλες εκμάθησης (learning curves) και ο error matrix.



Σχήμα 11: Καμπύλες εκμάθησης(Μοντέλο 3)



Σχήμα 12: Error Matrix (Μοντέλο 3)

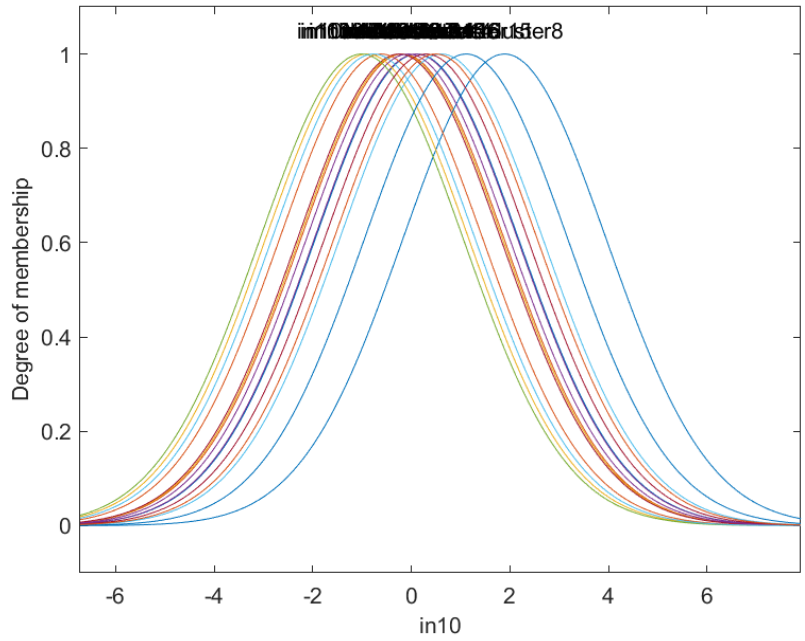
Στον παρακάτο πίνακα παρουσιάζονται οι μετρικές OA, PA, UA, $\hat{k}$.

Class	PA	UA	OA	$\hat{k}$
1	0.8836	0.1504	0.1580	0.0874
2	0.0041	1		
3	0.0012	0.0244		
4	0.0299	0.1844		
5	0.2367	0.3653		
6	0.1091	0.0472		
7	0.0775	0.0562		
8	0.0636	0.0337		
9	0.6824	0.3494		
10	0	0		
11	0.3562	0.1244		
12	0.2963	0.1509		

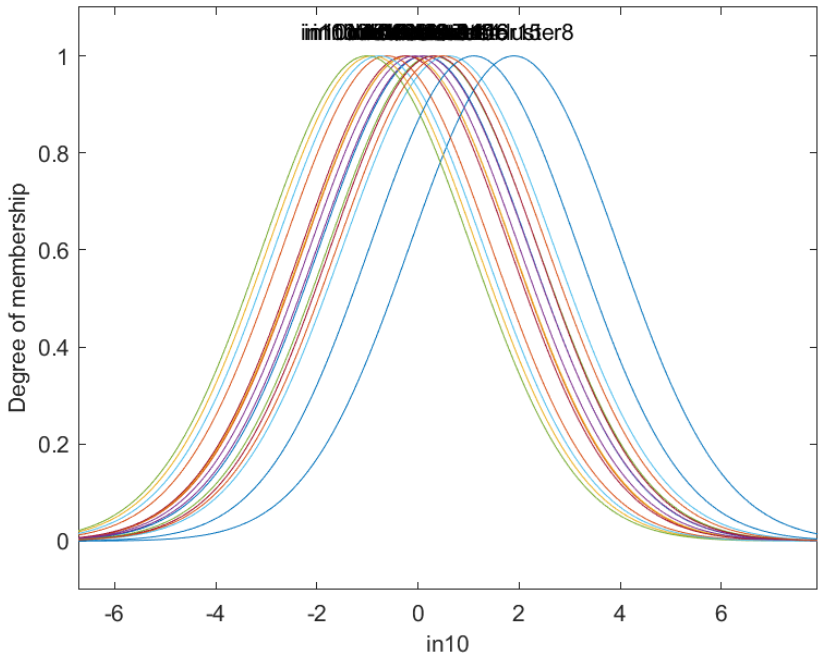
Πίνακας 7: Μετρικές OA, PA, UA, $\hat{k}$

Μοντέλο 4

Στα παρακάτω σχήματα παρουσιάζονται οι συναρτήσεις συμμετοχής πριν και μετά την εκπαίδευση.

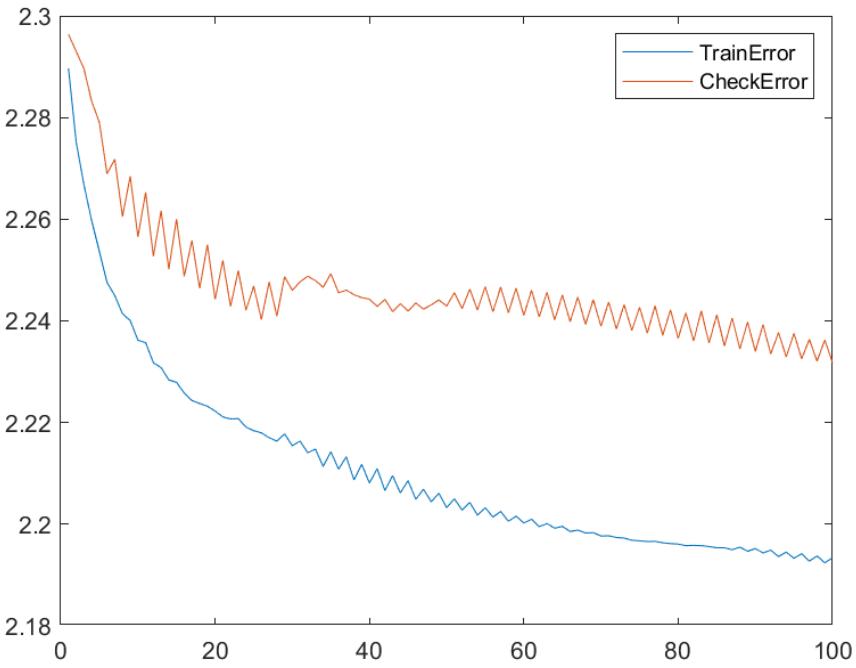


Σχήμα 13: Συναρτήσεις συμμετοχής πριν την εκπαίδευση (Μοντέλο 4)

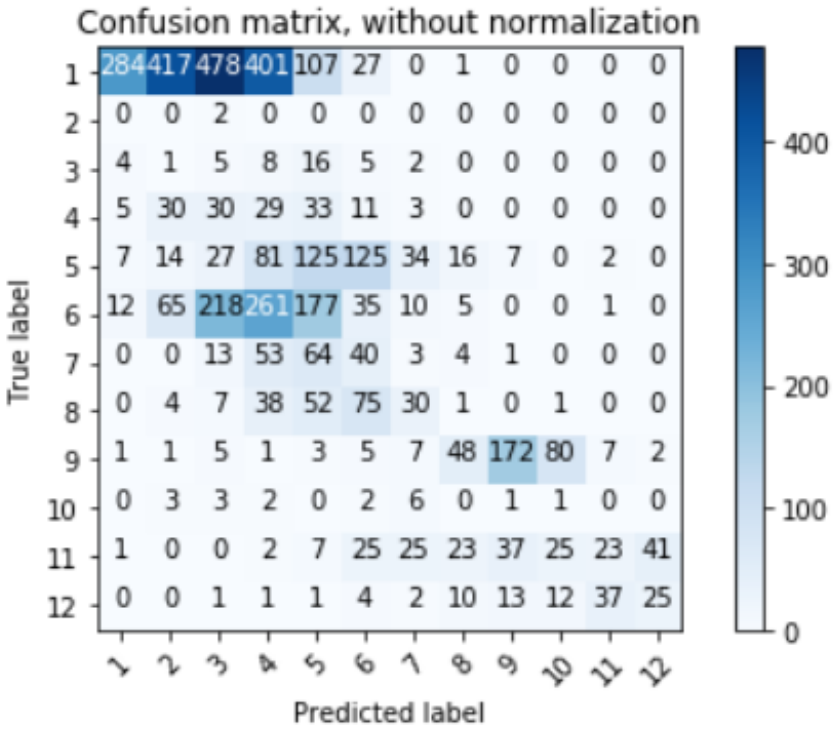


Σχήμα 14: Συναρτήσεις συμμετοχής μετά την εκπαίδευση (Μοντέλο 4)

Στη συνέχεια, παρουσιάζονται οι καμπύλες εκμάθησης (learning curves) και ο error matrix.



Σχήμα 15: Καμπύλες εκμάθησης(Μοντέλο 4)



Σχήμα 16: Error Matrix (Μοντέλο 4)

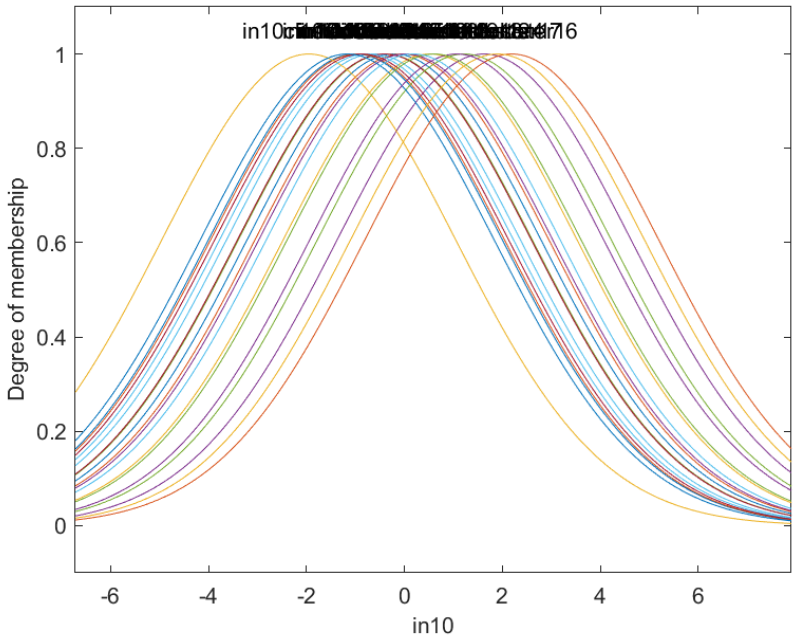
Στον παρακάτο πίνακα παρουσιάζονται οι μετρικές OA, PA, UA, $\hat{k}$.

Class	PA	UA	OA	$\hat{k}$
1	0.9045	0.1656	0.1685	0.0972
2	0	0		
3	0.0063	0.1220		
4	0.0331	0.2057		
5	0.2137	0.2854		
6	0.0989	0.0446		
7	0.0246	0.0169		
8	0.0093	0.0048		
9	0.7446	0.5181		
10	0.0084	0.0556		
11	0.3286	0.1100		
12	0.3676	0.2358		

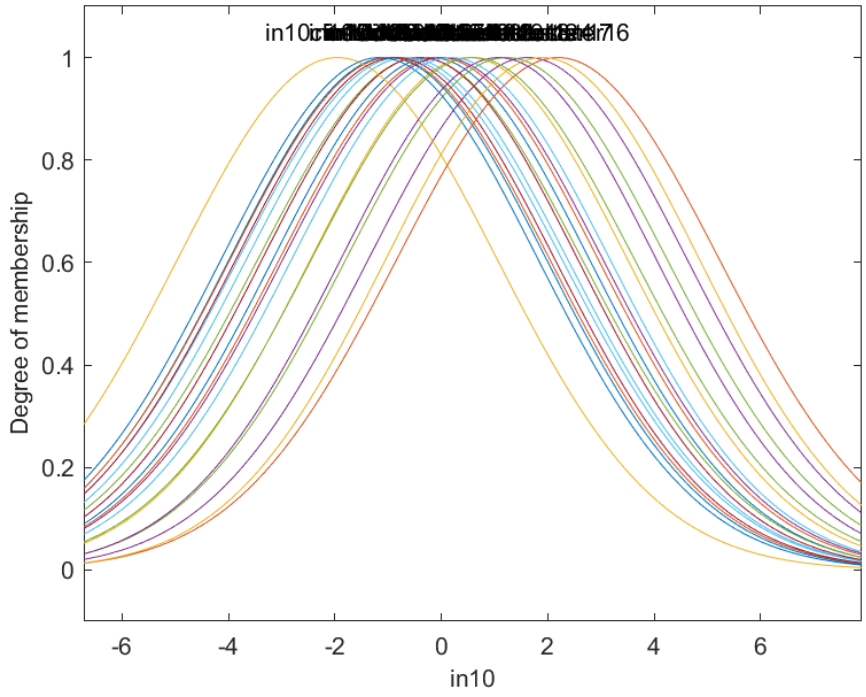
Πίνακας 8: Μετρικές OA, PA, UA, $\hat{k}$

Μοντέλο 5

Στα παρακάτω σχήματα παρουσιάζονται οι συναρτήσεις συμμετοχής πριν και μετά την εκπαίδευση.

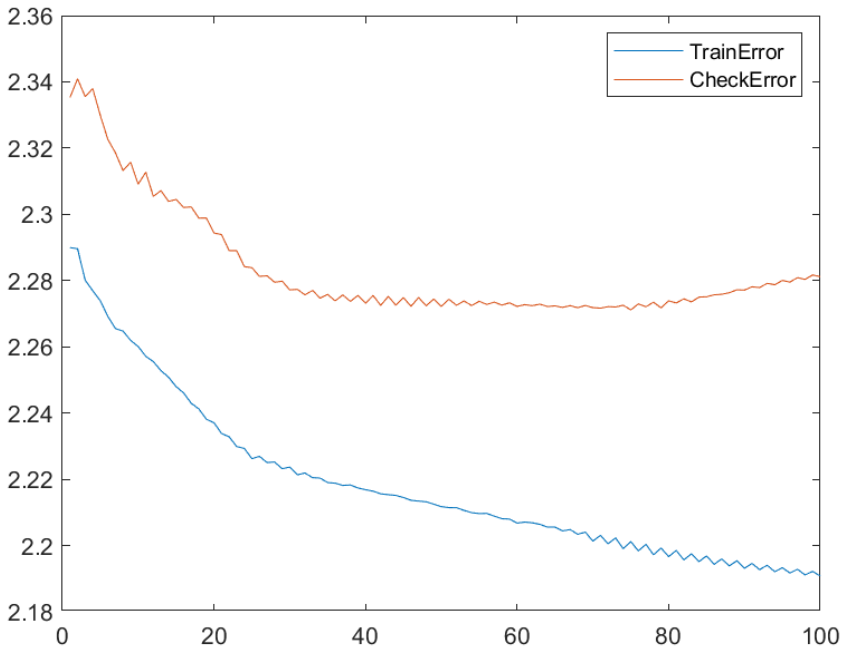


Σχήμα 17: Συναρτήσεις συμμετοχής πριν την εκπαίδευση (Μοντέλο 5)

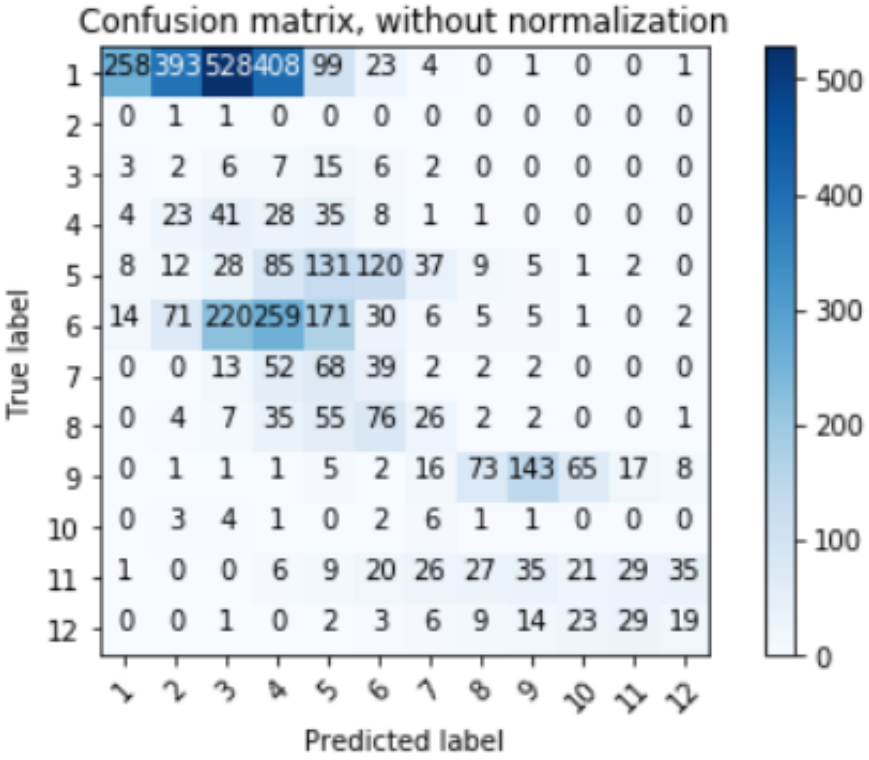


Σχήμα 18: Συναρτήσεις συμμετοχής μετά την εκπαίδευση (Μοντέλο 5)

Στη συνέχεια, παρουσιάζονται οι καμπύλες εκμάθησης (learning curves) και ο error matrix.



Σχήμα 19: Καμπύλες εκμάθησης(Μοντέλο 5)



Σχήμα 20: Error Matrix (Μοντέλο 5)

Στον παρακάτω πίνακα παρουσιάζονται οι μετρικές OA, PA, UA, $\hat{k}$.

Class	PA	UA	OA	$\hat{k}$
1	0.8958	0.1504	0.1556	0.0866
2	0.0020	0.5000		
3	0.0071	0.1463		
4	0.0317	0.1986		
5	0.2220	0.2991		
6	0.0912	0.0383		
7	0.0152	0.0112		
8	0.0155	0.0096		
9	0.6875	0.4307		
10	0	0		
11	0.3766	0.1388		
12	0.2879	0.1792		

Πίνακας 9: Μετρικές OA, PA, UA, $\hat{k}$

Σχολιασμός αποτελεσμάτων

Από τα παραπάνω αποτελέσματα βλέπουμε ότι κανένα από τα πέντε μοντέλα δεν έχει καλή απόδοση. Καλύτερες επιδόσεις μεταξύ τους έχει το μοντέλο 4 (εκείνο που έχει 16 κανόνες), με Overall Accuracy ίσο με 16,85% και $\hat{k}$ ίσο με 0.0972.

Ο βασικότερος λόγος που οφείλεται αυτή η χαμηλή απόδοση είναι το σετ δεδομένων. Πιο συγκεκριμένα, στο σετ δεδομένων, όπως μπορούμε να δούμε στον Πίνακα 1, δεν υπαρχεί ισσοροπία μεταξύ των κλάσεων. Η κλάση 1 αποτελεί σχεδόν το μισό dataset, ενώ οι κλάσεις όπως η 2 και 10 έχουν ποσοστό μοκρότερο από το 1%.

Εφαρμογή σε dataset με υψηλή διαστασιμότητα

Στη δεύτερη φάση της εργασίας θα ακολουύηθεί μια πιο συστηματική προσέγγιση στο πρόβλημα της χρήσης ασαφων νευρωνικών μοντέλων σε προβλήματα ταξινόμησης. Για το σκοπό αυτό θα επιλεχθεί ένα dataset με υψηλότερο βαύμό διαστασιμότητας. Ένα προφανές πρόβλημα που αναχύπτει από την επιλογή αυτή, είναι η λεγόμενη "έχρηξη" του πλήθους των IF-THEN χανόνων (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.
Σε αυτό το μέρος της εργασίας, χρησιμοποιείται το isolet dataset. Το συγκεκριμένο σετ δεδομένων είναι αρκετά μεγαλύτερο από το avila, αφού αποτελείται από 7797 εγραφές, που η κάθε μία έχει 617 διαφορετκά χαρακτηριστηκά.
Σύμφωνα με τα παραπάνω, μία αρκετά προφανής προσέγγιση είναι η μείωση του πλήθους των εισόδων και του πλήθους των κανόνων.

Διαχωρισμός του dataset και μείωση διαστάσεων και κανόνων

Αρχικά, πραγματοποιούμε τον διαχωρισμό του dataset σε τρία μη επικαλυπτόμενα υποσύολα $\{D_{trn}, D_{val}, D_{chk}\}$ ως εξής:

- D_{trn} = 60% του αρχικού σετ (Σετ το οποίο θα χρησιμοποιηθεί κατά την εκπαίδευση του μοντέλου).
- D_{val} = 20% του αρχικού σετ (Σετ το οποίο θα χρησιμοποιηθεί για την επικύρωση και αποφυγή του φαινομένου της υπερεκπαίδευσης).
- D_{chk} = 20% του αρχικού σετ (Σετ που θα χρησιμοποιηθεί για τον έλεγχο της απόδοσης του τελικού μοντέλου).

Για να επιτύχουμε καλή απόδοση, θα πρέπει η συχνότητα εμφάνισης δειγμάτων που ανήκουν σε μία συγκεκριμένη κλάση, σε κάθε ένα από τα τρία σύνολα διαμέρισης, να είναι όσο το δυνατόν πιο «όμοια» με την ανίστοιχη συχνότητα εμφάνισης τους στο αρχικό σύνολο δεδομένων. Πιο συγκεκριμένα, στους παρακάτω πίνακες φαίνεται η συχνότητα της κάθε κλάσης στο κάθε σύνολο.

Συχνότητα κάθε κλάσεις στο Συνολικό Dataset:

Class	Count	Percent
1	300	3.85%
2	300	3.85%
3	300	3.85%
4	300	3.85%
5	300	3.85%
6	298	3.82%
7	300	3.85%
8	300	3.85%
9	300	3.85%
10	300	3.85%
11	300	3.85%
12	300	3.85%
13	299	3.83%
14	300	3.85%
15	300	3.85%
16	300	3.85%
17	300	3.85%
18	300	3.85%
19	300	3.85%
20	300	3.85%
21	300	3.85%
22	300	3.85%
23	300	3.85%
24	300	3.85%
25	300	3.85%
26	300	3.85%

Πίνακας 10: συχνότητα κλάσεων στο συνολικό Dataset

Συχνότητα κάθε κλάσεις στο σετ εκπαίδευσης:

Class	Count	Percent
1	180	3.85%
2	180	3.85%
3	180	3.85%
4	180	3.85%
5	180	3.85%
6	179	3.83%
7	180	3.85%
8	180	3.85%
9	180	3.85%
10	180	3.85%
11	180	3.85%
12	180	3.85%
13	179	3.83%
14	180	3.85%
15	180	3.85%
16	180	3.85%
17	180	3.85%
18	180	3.85%
19	180	3.85%
20	180	3.85%
21	180	3.85%
22	180	3.85%
23	180	3.85%
24	180	3.85%
25	180	3.85%
26	180	3.85%

Πίνακας 11: συχνότητα κλάσεων στο σετ εκπαίδευσης

Συχνότητα κάθε κλάσεις στο σετ επικύρωσης:

Class	Count	Percent
1	60	3.85%
2	60	3.85%
3	60	3.85%
4	60	3.85%
5	60	3.85%
6	60	3.85%
7	60	3.85%
8	60	3.85%
9	60	3.85%
10	60	3.85%
11	60	3.85%
12	60	3.85%
13	60	3.85%
14	60	3.85%
15	60	3.85%
16	60	3.85%
17	60	3.85%
18	60	3.85%
19	60	3.85%
20	60	3.85%
21	60	3.85%
22	60	3.85%
23	60	3.85%
24	60	3.85%
25	60	3.85%
26	60	3.85%

Πίνακας 12: συχνότητα κλάσεων στο σετ επικύρωσης

Συχνότητα κάθε κλάσεις στο σετ ελέγχου:

Class	Count	Percent
1	60	3.85%
2	60	3.85%
3	60	3.85%
4	60	3.85%
5	60	3.85%
6	59	3.78%
7	60	3.85%
8	60	3.85%
9	60	3.85%
10	60	3.85%
11	60	3.85%
12	60	3.85%
13	60	3.85%
14	60	3.85%
15	60	3.85%
16	60	3.85%
17	60	3.85%
18	60	3.85%
19	60	3.85%
20	60	3.85%
21	60	3.85%
22	60	3.85%
23	60	3.85%
24	60	3.85%
25	60	3.85%
26	60	3.85%

Πίνακας 13: συχνότητα κλάσεων στο σετ ελέγχου

Στη συνέχεια, προκειμένου να μειώσουμε τις διαστάσεις της κάθε εγγραφής του σετ δεδομένων, εφαρμόζουμε τον αλγόριθμο Relief. Ο παραπάνω αλγόριθμος επιστρέφει τα πιο σημαντικά χαρακτηριστηκά (απο το πιο σημαντικό στο λιγότερο).

Έπειτα, προκειμένου να βρούμε το χαρακτηριστηκά του μοντέλου που ανταποκρίνεται καλύτερα στο πρόβλημα εφαρμόζουμε την μέθοδο Grid Search σε συνδυασμό με τον K-fold Cross Validation (όπου k = 5).

Κατά την διαδικασία του Grid Search και του K-fold Cross Validation, αρχικά διαχωρίζουμε το σετ εκπαίδευσης σε 5 διαφορετικά σετ, κρατώντας το ποσοστό εμφάνισης της κάθε κλάσης περίπου σταθερό. Σε κάθε μία από τις 5 επαναλήψεις (λόγω του k = 5), τα 4 από τα 5 μέρη του σετ χρησιμοποιούνται για την εκπαίδευση του μοντέλου και το $5^{ο}$ για την επικύρωση του. Στη συνέχεια, υπολογίζεται η μέση τιμή του μέσου τετραγωνικού σφάλματος (MSE), ώστε αργότερα να γίνει η σύγκριση των διαφορετικών μοντέλων (διαφορετικά μοντέλα θεωρούνται εκείνα που έχουν διαφορετικό αριθμό κανόνων και χρισιμοποιούν διαφορετικό αριθμό χαρακτηριστηκών του dataset) και να επιλεχθεί το καλύτερο.

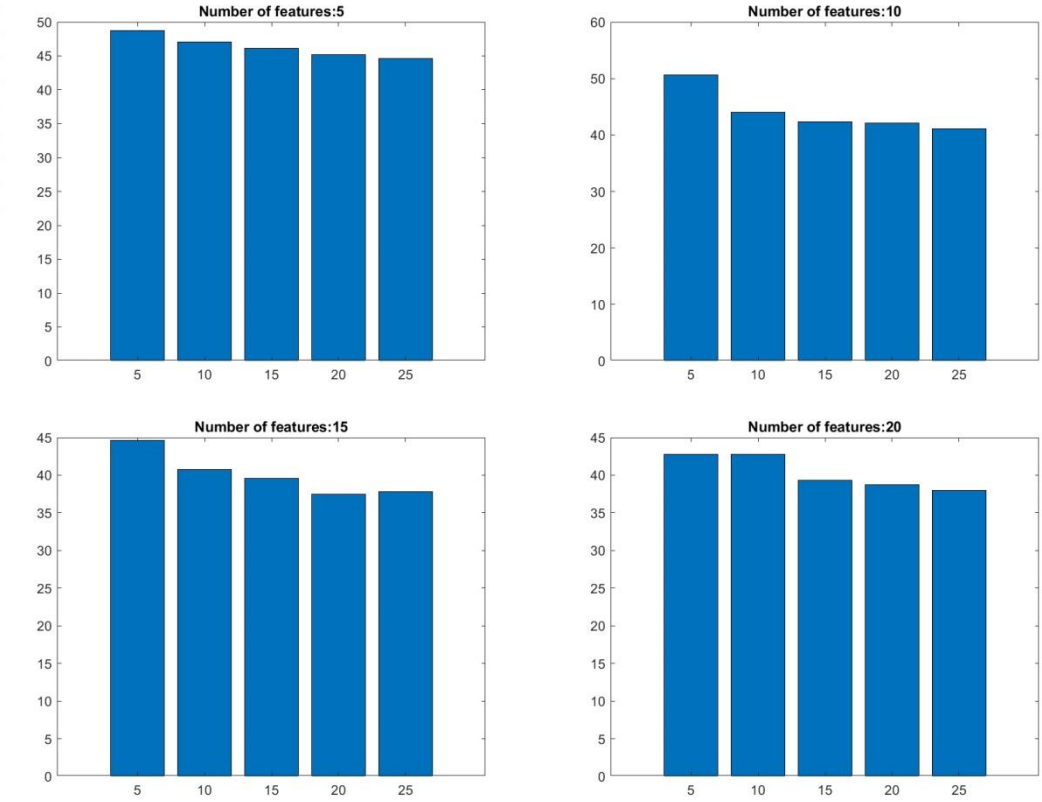
Για την ομαδοποίηση και τη δημιουργία των κανόνων, χρησιμοποιήθηκε η μέθοδος FCM (Fuzzy C-Means). Τα μοντέλα που συγκρίθηκαν οποτελούνται από το συνδυασμό χαρακτηριστικών και κανόνων που προκύπτουν από το καρτεσιανό γινόμενο:
NF×NR = {5, 10, 15, 20} × {5, 10, 15, 20, 25}

Στον παρακάτω πίνακα παρουσιάζεται η μέση τιμή του MSE των μοντέλων.

Rules / Feature	5	10	15	20	25
5	48.7028	47.0994	46.1436	45.2167	44.5983
10	50.5868	44.0021	42.3778	42.1226	41.0733
15	44.6107	40.7562	39.5679	37.4707	37.7665
20	42.7959	42.7748	39.3575	38.7150	37.9793

Πίνακας 14: Μέση τιμή MSE μοντέλων

Στο παρακάτω σχήμα παρουσιάζονται η μέση τιμή των MSE σε γραφική μορφή.

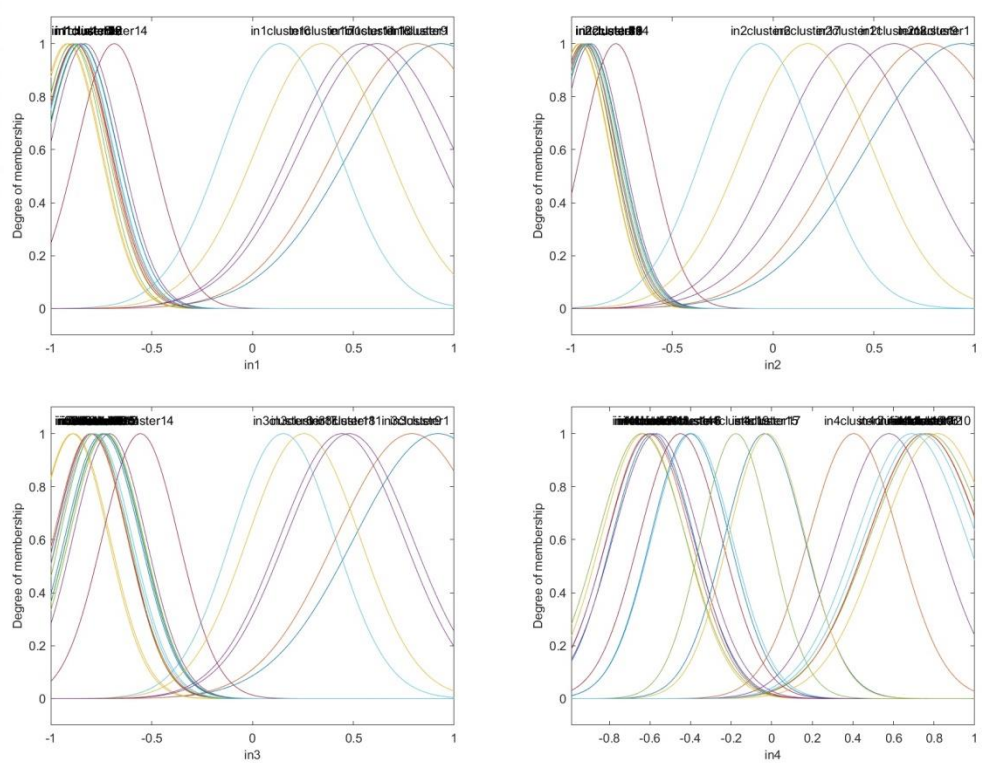


Σχήμα 21: Ιστόγραμμα των MSE των μοντέλων

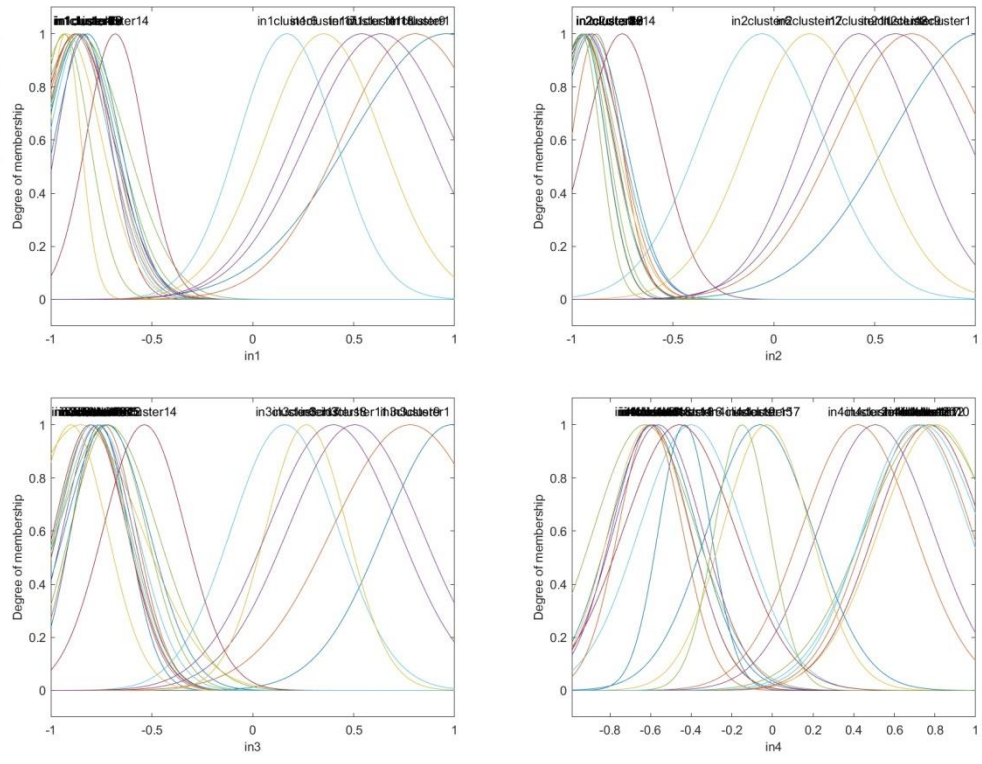
Βέλτιστο Μοντέλων (15 features, 20 rules)

Από τον πίνακα 14, βλέπουμε ότι το βέλτιστο μοντέλο είναι εκείνο που εκπαιδεύεται με 15 χαρακτηριστικά, από τα 617, και έχει 20 IF THEN κανόνες.

Στα παρακάτω διαγράμματα παρουσιάζονται ενδεικτικά μερικά ασαφή σύνολα στην αρχική και τελική τους μορφή.

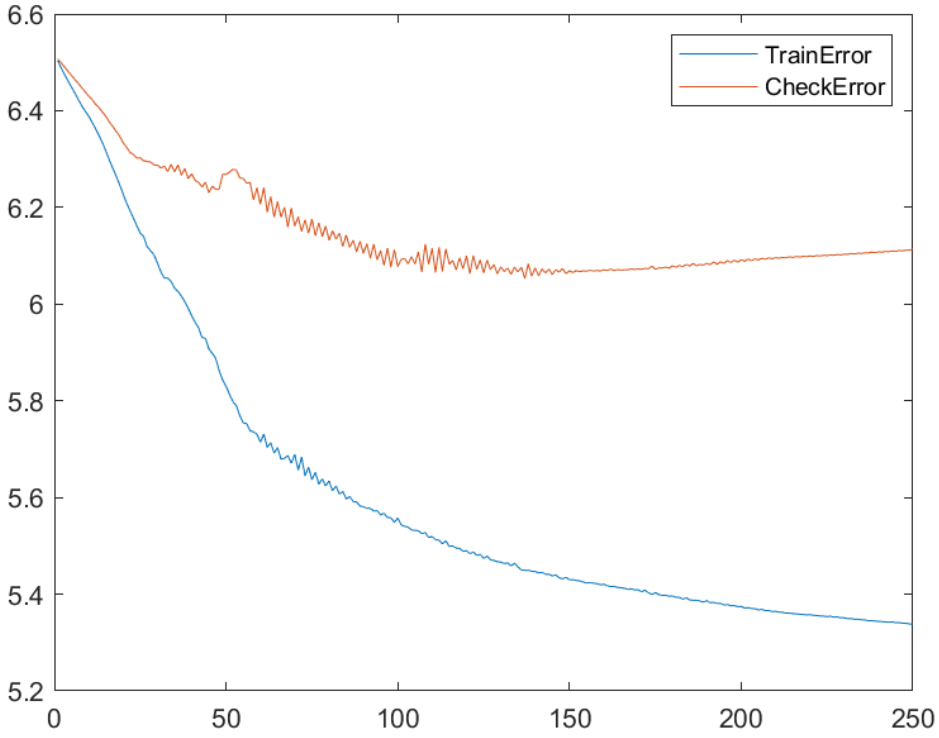


Σχήμα 22: Ενδεικτικές συναρτήσεις συμμετοχης πριν την εκπαίδευση

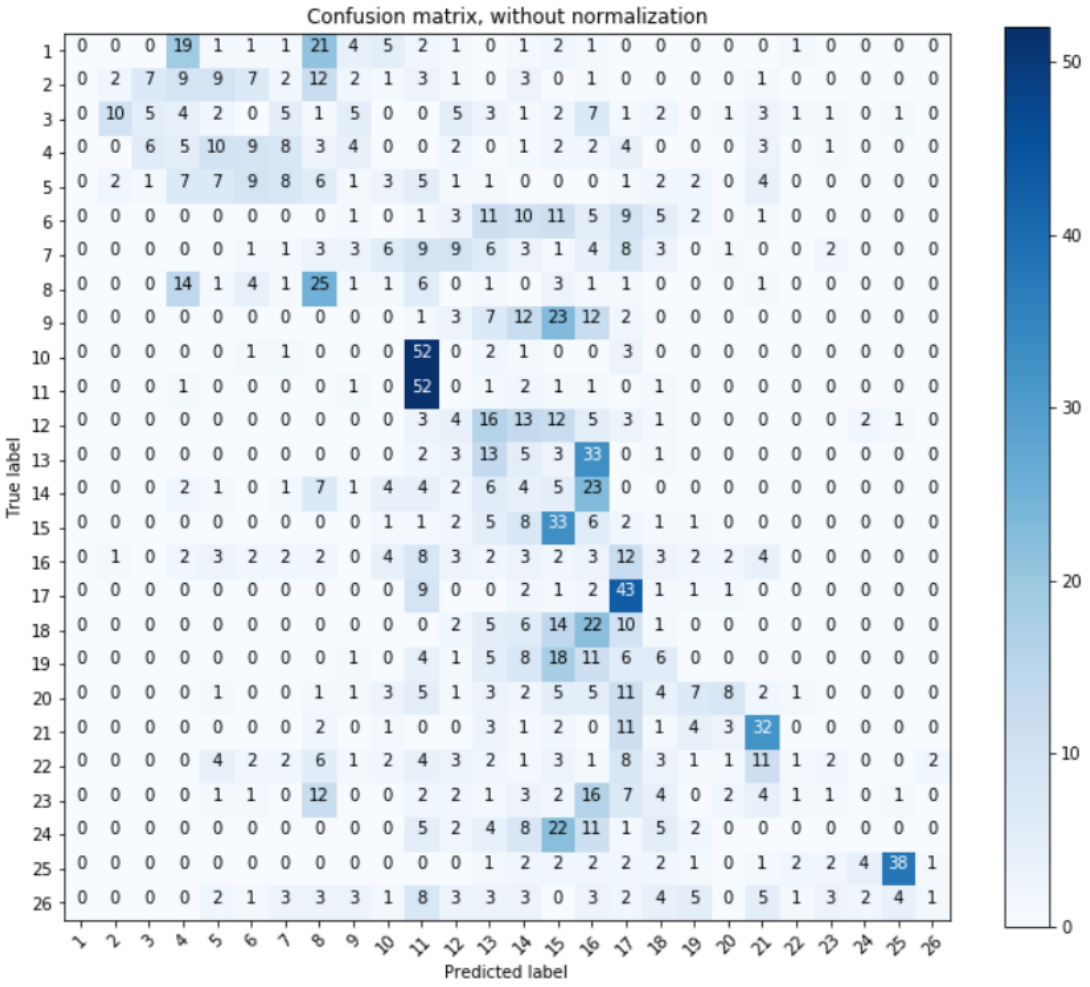


Σχήμα 23: Ενδεικτικές συναρτήσεις συμμετοχης μετά την εκπαίδευση

Στη συνέχεια, παρουσιάζονται οι καμπύλες εκμάθησης (learning curves) και ο error matrix.



Σχήμα 24: Καμπύλες εκμάθησης



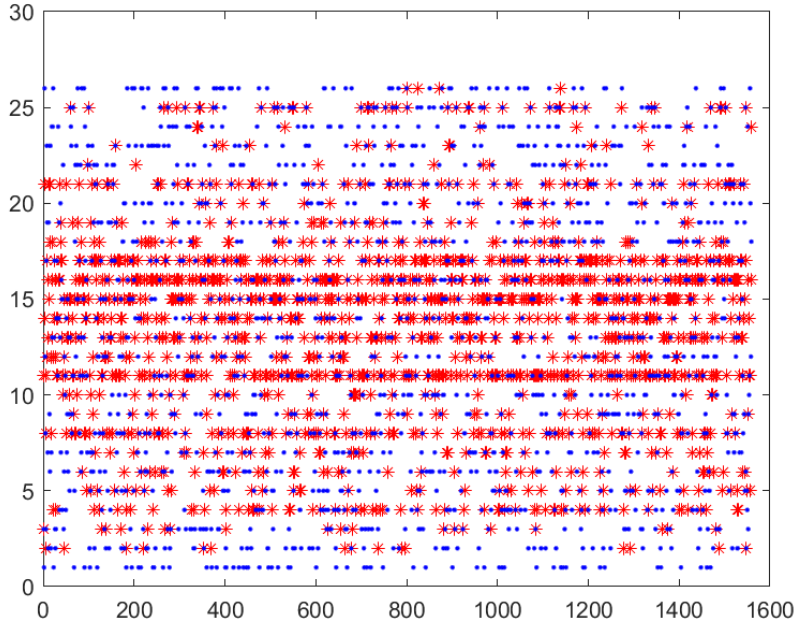
Σχήμα 25: Error Matrix

Στον παρακάτω πίνακα παρουσιάζονται οι μετρικές OA, PA, UA, $\hat{k}$.

Class	PA	UA	Class	PA	UA	Accuracy
1	NaN	0	14	0.0388	0.0667	17.90%
2	0.1333	0.0333	15	0.1953	0.5500	
3	0.2632	0.0833	16	0.0169	0.0500	
4	0.0794	0.0833	17	0.2925	0.7167	
5	0.1667	0.1167	18	0.0200	0.0167	
6	0	0	19	0	0	$\hat{k}$
7	0.0286	0.0167	20	0.4211	0.1333	
8	0.2404	0.4167	21	0.4444	0.5333	0.1461
9	0	0	22	0.1250	0.0167	
10	0	0	23	0.0833	0.0167	
11	0.2796	0.8667	24	0	0	
12	0.0755	0.0667	25	0.8444	0.6333	
13	0.1287	0.2167	26	0.2500	0.0167	

Πίνακας 10: Μετρικές OA, PA, UA, $\hat{k}$

Τέλος, στο παρακάτω διάγραμμα παρουσιάζονται οι προβλέψεις του τελικού μοντέλου καθώς και οι πραγματικές τιμές.



Σχήμα 25: Error Matrix

Με κόκκινο είναι οι προβλέψεις του μοντέλου και με μπλέ είναι οι πραγματικές τιμές.

Σχολιασμός αποτελεσμάτων

Από το παραπάνω πίνακα (Πίνακας 10), βλέπουμε ότι το μοντέλο, μετά από 250 εποχές εκπαίδευσης, πετυχαίνει accuracy ίσο με 17.9% και $\hat{k}$ ίσο με 0.1461. Τα αποτελέσματα του μοντέλου δεν είναι και πολύ ικανοποιητικά. Το γεγονός αυτό πιθανότατα οφείλεται στο μικρό μέγεθος του σετ δεδομένων, αν αναλογιστούμε ότι το σετ έχει αρκετά μεγάλο πλήθος κλάσεων (26 διαφορετικές κλάσεις). Επίσης, από τις καμπύλες εκμάθησης βλέπουμε οτι το μοντέλο είναι επιρεπές όσον αφορά την υπερεκπαίδευση, οπότε η χρήση περισσότερων εποχών κατά την εκπαίδευση δεν θα βελτίωνε τα αποτελέσματα.

Αρχεία

- avila.m: Αρχείο εκπαίδευσης 5 μοντέλων με διαφορετικά χαρακτηριστικά για το dataset Avila.
- plotmfin.m: Συνάρτηση για τη δημιουργία και την αποθήκευση των συναρτήσεων συμμετοχης εισόδου.
- gridSearch.m: Αρχείο επιλογής του καταλληλότερου μοντέλου για το dataset Isolet.
- optimum_model: Εκπαίδευση και αξιολόγηση του βέλτιστου μοντέλου για το dataset Isolet.
- plotmfin.m: Συνάρτηση για τη δημιουργία και την αποθήκευση των συναρτήσεων συμμετοχης εισόδου.