Τμήμα Μηχανικών Η/Υ & Πληροφορικής

Πανεπιστήμιο Ιωαννίνων

Μεταπτυχιακό μάθημα: T06 - Μηχανική Μάθηση

Σκορδά Ελένη Α.Μ.: 374

Δημητριάδης Σωκράτης Α.Μ.: 359

**Εργασία 1**

*Άσκηση 1*

Α) Ο ταξινομητής των Κ-κοντινότερων γειτόνων, χρησιμοποιώντας την Ευκλείδεια απόσταση και έναν έως και δέκα γείτονες, κατηγοριοποίησε τα δεδομένα-χαρακτήρες δίνοντας τα εξής αποτελέσματα:

Πλήθος Γειτόνων : 1

Απώλεια Μεθόδου = 0.0217

Ποσοστό Επιτυχίας = 0.9638

Πλήθος Γειτόνων : 2

Απώλεια Μεθόδου = 0.0256

Ποσοστό Επιτυχίας = 0.9633

Πλήθος Γειτόνων : 3

Απώλεια Μεθόδου = 0.0207

Ποσοστό Επιτυχίας = 0.9649

Πλήθος Γειτόνων : 4

Απώλεια Μεθόδου = 0.0243

Ποσοστό Επιτυχίας = 0.9644

Πλήθος Γειτόνων : 5

Απώλεια Μεθόδου = 0.0262

Ποσοστό Επιτυχίας = 0.9649

Πλήθος Γειτόνων : 6

Απώλεια Μεθόδου = 0.0269

Ποσοστό Επιτυχίας = 0.9638

Πλήθος Γειτόνων : 7

Απώλεια Μεθόδου = 0.0246

Ποσοστό Επιτυχίας = 0.9638

Πλήθος Γειτόνων : 8

Απώλεια Μεθόδου = 0.0235

Ποσοστό Επιτυχίας = 0.9638

Πλήθος Γειτόνων : 9

Απώλεια Μεθόδου = 0.0264

Ποσοστό Επιτυχίας = 0.9622

Πλήθος Γειτόνων : 10

Απώλεια Μεθόδου = 0.0269

Ποσοστό Επιτυχίας = 0.9610

Την καλύτερη απόδοση, παρατηρούμε ότι την δίνει το μοντέλο των τριών γειτόνων, καθώς έχει το μέγιστο ποσοστό επιτυχίας (μαζί με αυτό των πέντε γειτόνων) και τη μικρότερη απώλεια.

Ο αντίστοιχος ταξινομητής των Κ-κοντινότερων γειτόνων που χρησιμοποιεί την απόσταση Hamming και έναν έως και δέκα γείτονες, κατηγοριοποίησε τα δεδομένα-χαρακτήρες, δίνοντας τα εξής αποτελέσματα:

Πλήθος Γειτόνων : 1

Απώλεια Μεθόδου = 0.1138

Ποσοστό Επιτυχίας = 0.8581

Πλήθος Γειτόνων : 2

Απώλεια Μεθόδου = 0.1381

Ποσοστό Επιτυχίας = 0.8525

Πλήθος Γειτόνων : 3

Απώλεια Μεθόδου = 0.1067

Ποσοστό Επιτυχίας = 0.8742

Πλήθος Γειτόνων : 4

Απώλεια Μεθόδου = 0.1104

Ποσοστό Επιτυχίας = 0.8826

Πλήθος Γειτόνων : 5

Απώλεια Μεθόδου = 0.1054

Ποσοστό Επιτυχίας = 0.8792

Πλήθος Γειτόνων : 6

Απώλεια Μεθόδου = 0.1031

Ποσοστό Επιτυχίας = 0.8798

Πλήθος Γειτόνων : 7

Απώλεια Μεθόδου = 0.1020

Ποσοστό Επιτυχίας = 0.8792

Πλήθος Γειτόνων : 8

Απώλεια Μεθόδου = 0.1031

Ποσοστό Επιτυχίας = 0.8770

Πλήθος Γειτόνων : 9

Απώλεια Μεθόδου = 0.0999

Ποσοστό Επιτυχίας = 0.8770

Πλήθος Γειτόνων : 10

Απώλεια Μεθόδου = 0.1020

Ποσοστό Επιτυχίας = 0.8776

Παρατηρείται ότι με τέσσερις γείτονες έχουμε μεγαλύτερο ποσοστό επιτυχίας αλλά μεγαλύτερη απώλεια από ότι άλλες με περισσότερους γείτονες.

Εν κατακλείδει, η Ευκλείδεια απόσταση δίνει εμφανώς καλύτερα αποτελέσματα στην ταξινόμηση από την Hamming.

Γ) Στην ταξινόμηση με πολυεπίπεδο νευρωνικό δίκτυο, έγινε χρήση του εργαλείου της Matlab 2016b nprtool όπου εκτελέσαμε δύο πειράματα με τις εξής ρυθμίσεις:

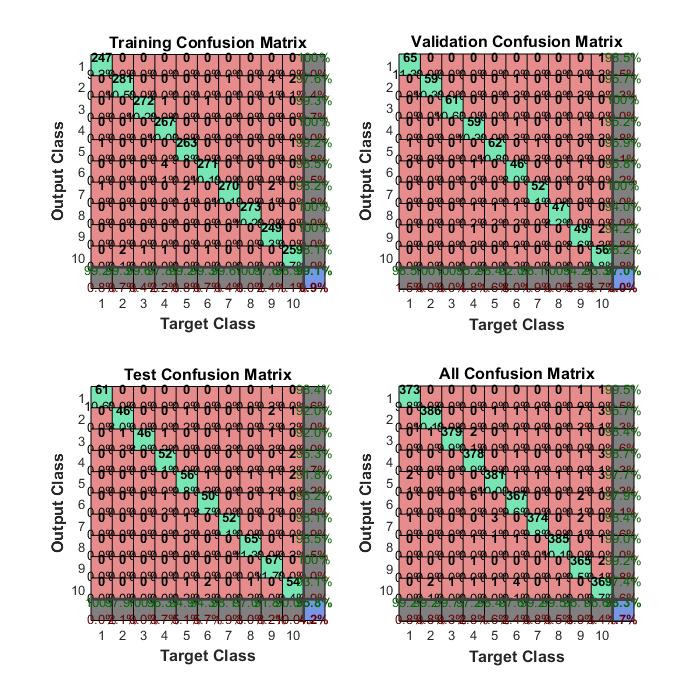
1η δοκιμή:

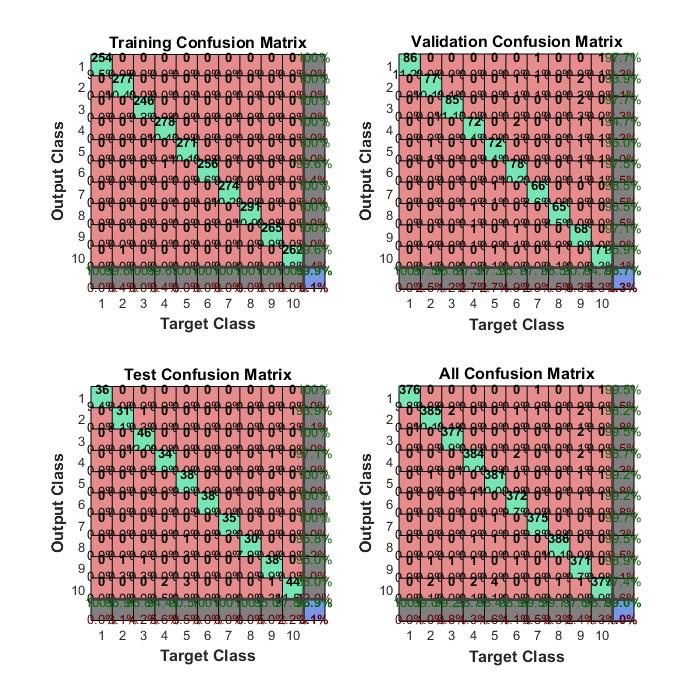
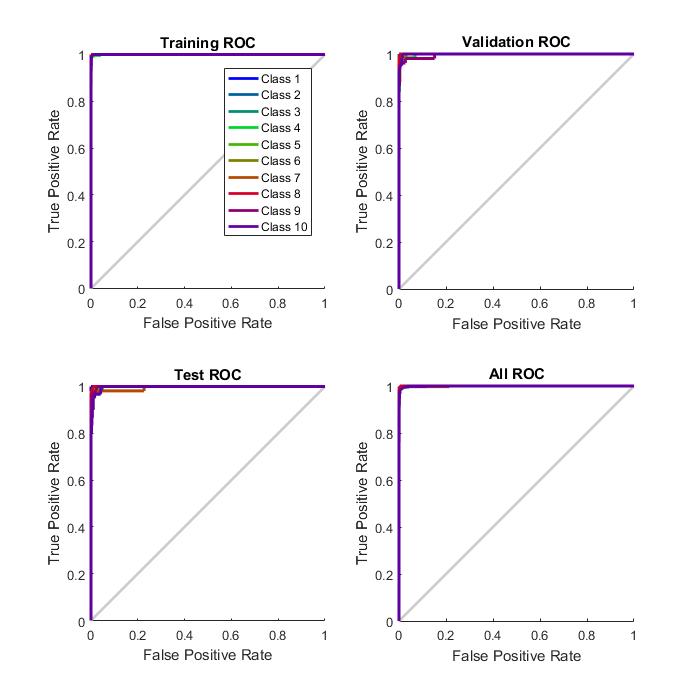
Επιλέχθηκε το 70% των δειγμάτων για εκπαίδευση (2677 δείγματα), 15% για εξακρίβωση (573 δείγματα) και 15% για έλεγχο (573 δείγματα). Επίσης τέθηκαν σε λειτουργία 20 νευρώνες με ένα κρυμμένο επίπεδο για τον σκοπό της ταξινόμησης.

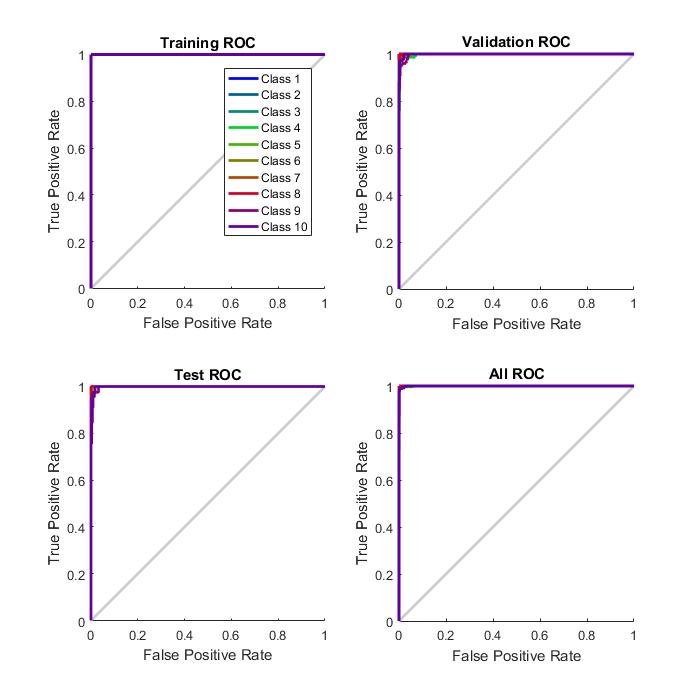
2η δοκιμή:

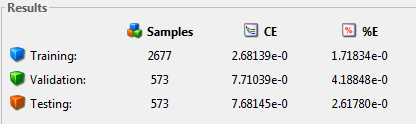
Επιλέχθηκε το 70% των δειγμάτων για εκπαίδευση (2676 δείγματα), 20% για εξακρίβωση (765 δείγματα) και 10% για έλεγχο (382 δείγματα). Επίσης τέθηκαν σε λειτουργία 30 νευρώνες με ένα κρυμμένο επίπεδο για τον σκοπό της ταξινόμησης.

Οι γραφικές παραστάσεις που προέκυψαν είναι αντίστοιχα οι εξής:

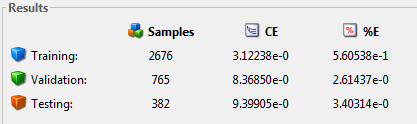






Οι μετρήσεις ήταν αντίστοιχα:  


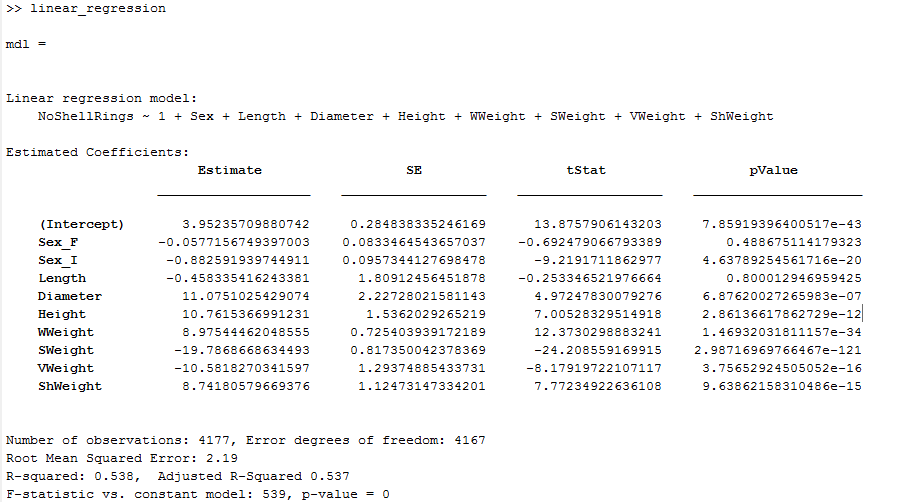
και



Παρατηρείται ότι η δεύτερη δοκιμή έχει μικρότερα ποσοστά σφαλμάτων στην εκπαίδευση και την εξακρίβωση, αλλά μεγαλύτερο στον έλεγχο από την πρώτη.

*Άσκηση 2:*

Α) Τα αποτελέσματα της Γραμμικής Παλινδρόμησης είναι:



Δ) Στην MLP με ένα κρυμμένο επίπεδο, xρησιμοποιήθηκαν 30 νευρώνες. Τα αποτελέσματα ήταν τα εξής:

