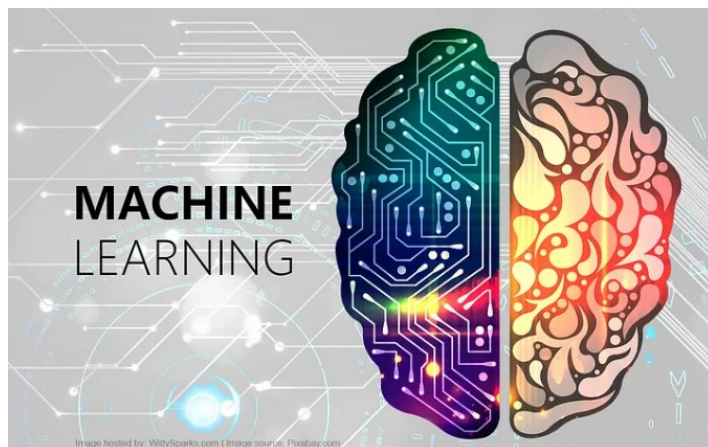




Τμήμα Μηχανικών Η/Υ & Πληροφορικής,
Πανεπιστήμιο Ιωαννίνων,
Ακαδημαϊκό έτος 2023-24

Προπτυχιακό μάθημα: Μηχανική Μάθηση
1η Σειρά Ασκήσεων



Ονόματα ομάδας, ΑΜ: Βασίλειος Παπακυριάκου, 5324
Δανάη Χανλαρίδου, 5386

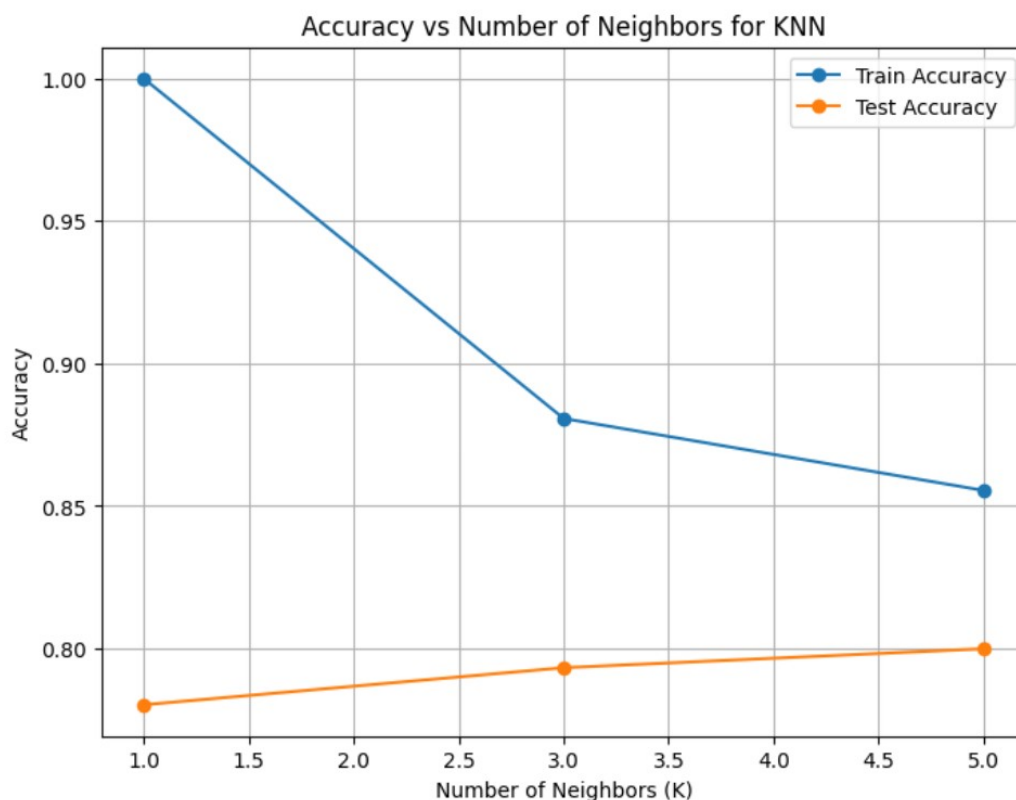
Ημερομηνία παράδοσης: 14/05/2024

Σχολιασμός για επίδοση και συμπεριφορά των μοντέλων

[A]. (διανυσματική αναπαράσταση)

- **Nearest Neighbor Classifier**

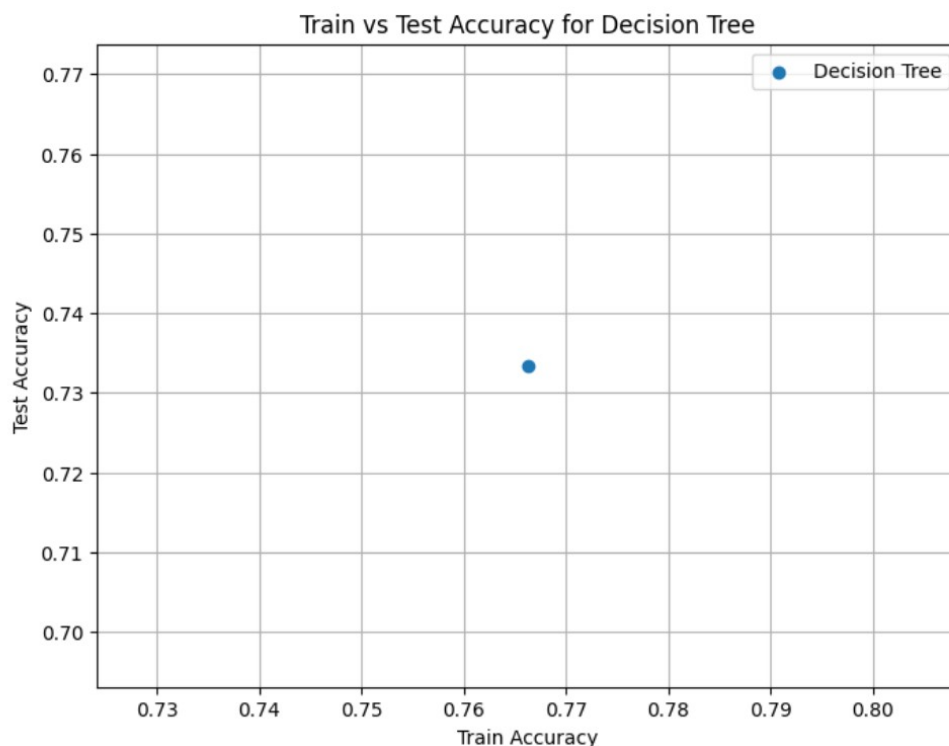
Η επίδοση του Nearest Neighbor Classifier στο Fashion-MNIST dataset μπορεί να είναι ικανοποιητική, αλλά εξαρτάται σημαντικά από τη συγκεκριμένη διαμόρφωση του προβλήματος και τις επιλογές των υπερπαραμέτρων. Από το διάγραμμα, παρατηρούμε πως όσο αυξάνεται ο αριθμός K γειτόνων, τόσο καλύτερη ακρίβεια έχουμε.



Σχήμα 1

- **Decision Tree**

Η επίδοση του Decision Tree εξαρτάται σε μεγάλο βαθμό από τη σωστή επιλογή των υπερπαραμέτρων, όπως το μέγιστο βάθος (max depth) του δέντρου. Εάν, το μέγιστο βάθος είναι πολύ μεγάλο, το δέντρο μπορεί να υπερπροσαρμοστεί στα δεδομένα εκπαίδευσης και να έχει κακή γενίκευση στα δεδομένα ελέγχου. Από την άλλη πλευρά, εάν το μέγιστο βάθος είναι πολύ μικρή, το δέντρο μπορεί να μην είναι ικανό να μάθει την πολυπλοκότητα των δεδομένων και να έχει υπο-επίδοση. Επίσης, η οπτικοποίηση του δέντρου μπορεί να βοηθήσει στην κατανόηση των αποφάσεων που λαμβάνει το μοντέλο. Από το διάγραμμα, παρατηρούμε ότι η ακρίβεια της εκπαίδευσης είναι περίπου 0.765 ενώ η ακρίβεια του ελέγχου είναι περίπου 0.735.

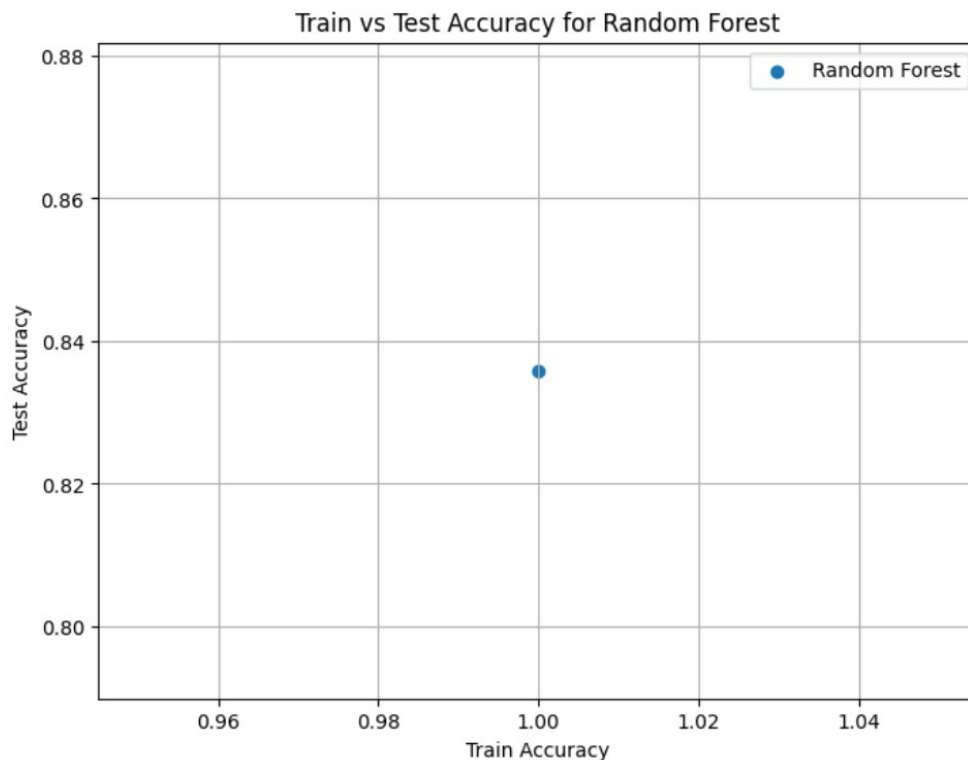


Σχήμα 2

- **Random Forest (100 estimators)**

Ο Random Forest μπορεί να προσφέρει υψηλή ακρίβεια λόγω της πολυπλοκότητας των δεδομένων. Ωστόσο, η επίδοση του εξαρτάται από τις υπερπαραμέτρους του, όπως ο αριθμός των δέντρων και η μέγιστη βάθμωση (max depth) των δέντρων. Συνολικά, ο Random Forest είναι ένα ισχυρό μοντέλο που μπορεί να παράσχει αξιόπιστες προβλέψεις για το Fashion-MNIST dataset. Από το διάγραμμα, μπορούμε να παρατηρήσουμε ότι η ακρίβεια της εκπαίδευσης είναι 1

ακριβώς ενώ η ακρίβεια του ελέγχου είναι περίπου 0.84, δηλαδή μια αρκετά καλή πρόσεγγιση.



Σχήμα 3

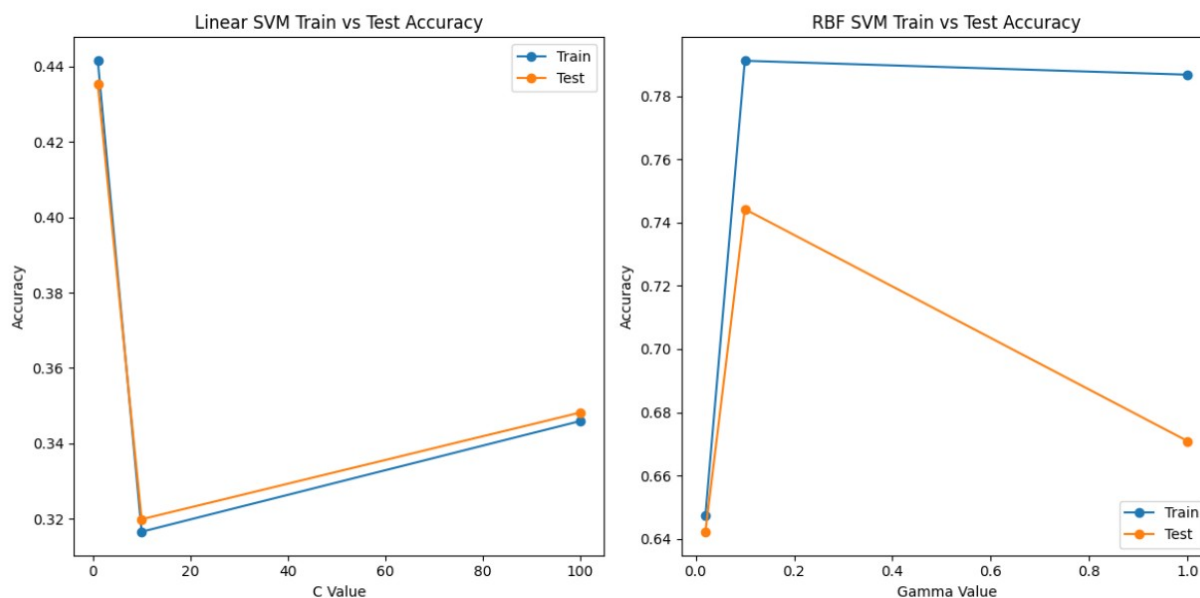
- **Linear SVM classifier ($C = \{1, 10, 100\}$)**

Για το Fashion-MNIST dataset, ο γραμμικός ταξινομητής SVM με διάφορες τιμές της παραμέτρου C μπορεί να επιτύχει καλή ακρίβεια, αλλά είναι σημαντικό να επιλεγεί η βέλτιστη τιμή της παραμέτρου C μέσω διασταύρωσης για να αποφευχθεί η υπερ-ή υπο-προσαρμογή. Επιπλέον, ο γραμμικός ταξινομητής SVM είναι ιδιαίτερα αποδοτικός για μεγάλα σύνολα δεδομένων, καθιστώντας τον κατάλληλο για το Fashion-MNIST dataset. Από το πρώτο διάγραμμα του σχήματος 4 παρατηρούμε ότι για κάθε τιμή του C η ακρίβεια της εκπαίδευσης και του ελέγχου ταυτίζονται σε μεγάλο βαθμό.

- **SVM with kernel RBF ($\gamma = 0.02, 0.1, 1$)**

Ο SVM με πυρήνα RBF μπορεί να παρέχει υψηλή ακρίβεια στο Fashion-MNIST dataset λόγω της ικανότητάς του να αναγνωρίζει πολύπλοκα πρότυπα στα δεδομένα. Η απόδοσή του εξαρτάται σημαντικά από την υπερπαραμέτρο γ (gamma), η οποία ελέγχει την επίδραση του κάθε δείγματος στην απόφαση του μοντέλου. Από το διάγραμμα παρατηρούμε ότι όσο μεγαλώνει η παράμετρος γ τόσο

χειροτερεύει η ακρίβεια μεταξύ της εκπαίδευσης και του ελέγχου. Η καλύτερη τιμή της παραμέτρου γ φαίνεται να είναι η 0.02 κατά την οποία η ακρίβεια ελέγχου είναι περίπου 0.64 και η ακρίβεια εκπαίδευσης περίπου 0.65, ενώ για τις τιμές 0.1 και 1 υπάρχει μεγαλύτερη απόκλιση μεταξύ ακρίβεια ελέγχου και εκπαίδευσης.



Σχήμα 4

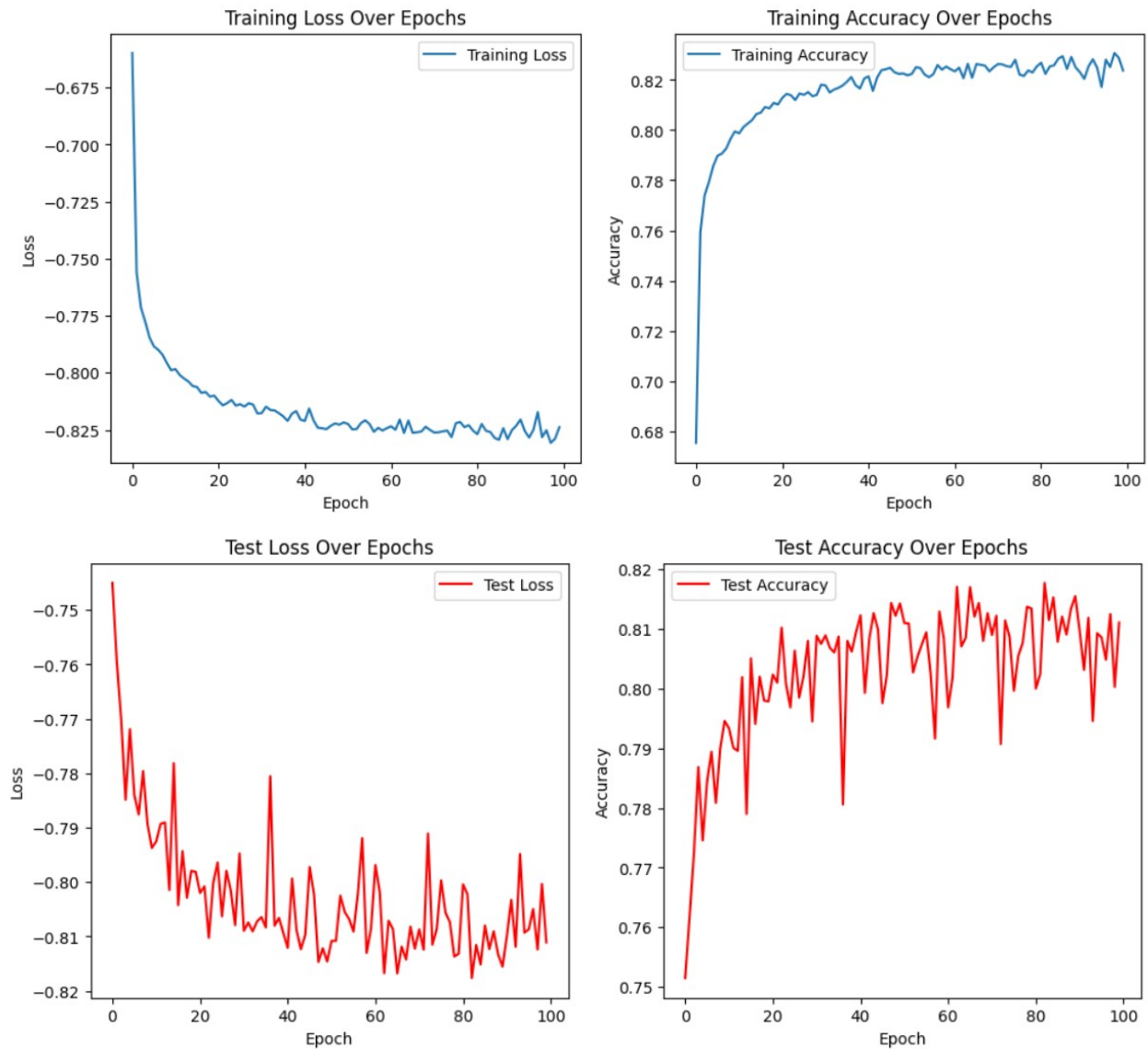
Ως επιλογή μεθόδου ταξινόμησης φαίνεται να είναι καλύτερη η πρώτη καθώς για όλες τις τιμές υπάρχει καλύτερη προσέγγιση.

- **Feed-forward Neural Network (3 hidden layers)**

Η συμπεριφορά του νευρωνικού δικτύου εξαρτάται σημαντικά από την αρχιτεκτονική του (πλήθος επιπέδων, πλήθος νευρώνων κλπ.) και τις υπερπαραμέτρους του (όπως ο ρυθμός μάθησης, η παράμετρος επαύξησης κλπ.). Η επιτυχία του νευρωνικού δικτύου εξαρτάται επίσης από τον τρόπο εκπαίδευσής του, την αποφυγή της υπερ-προσαρμογής και τη βελτιστοποίηση των υπερπαραμέτρων.

Εκπαιδεύουμε το δίκτυο με 100 εποχές και batch size 50. Παρακάτω φαίνονται τα διαγράμματα της συνάρτησης απώλειας και ακρίβειας και για τα δύο σύνολα (εκπαίδευσης και ελέγχου).

Παρατηρούμε πως καθώς αυξάνεται ο αριθμός των εποχών η τιμή της συνάρτησης απώλειας μειώνεται και η τιμή της ακρίβειας αυξάνεται και για τα δύο σύνολα.



Σχήμα 6

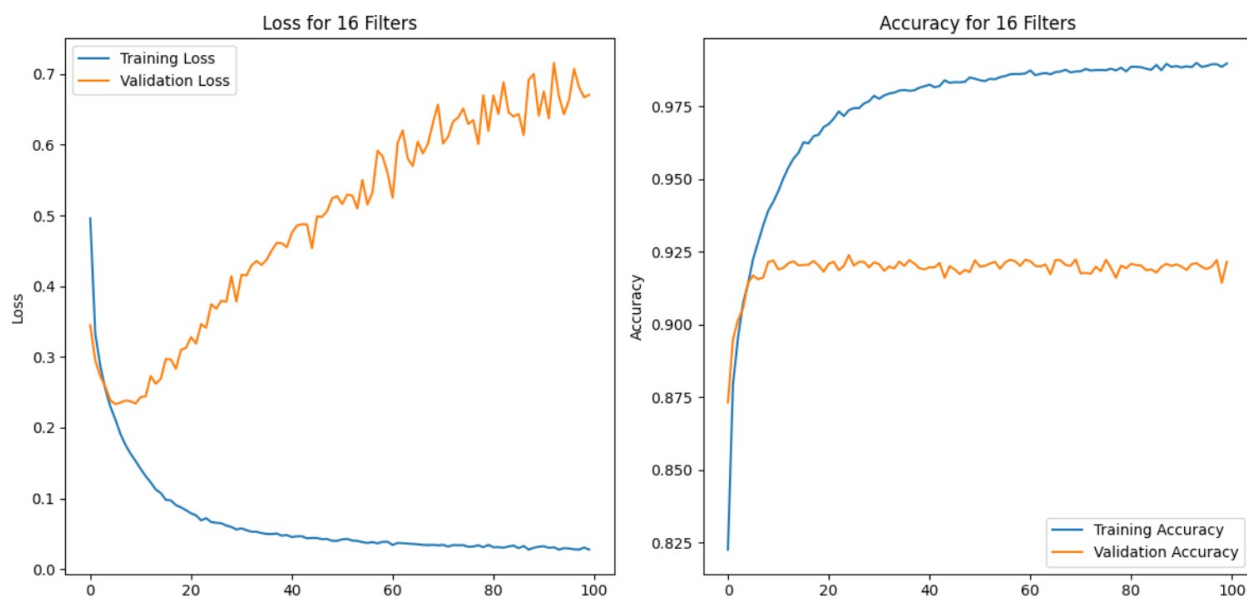
[B]. (δεδομένα με την μορφή εικόνας)

Ο αριθμός των φίλτρων σε κάθε συνελικτικό επίπεδο επηρεάζει την επίδοση του δικτύου στα δεδομένα του συνόλου ελέγχου (testing set) με τους ακόλουθους τρόπους:

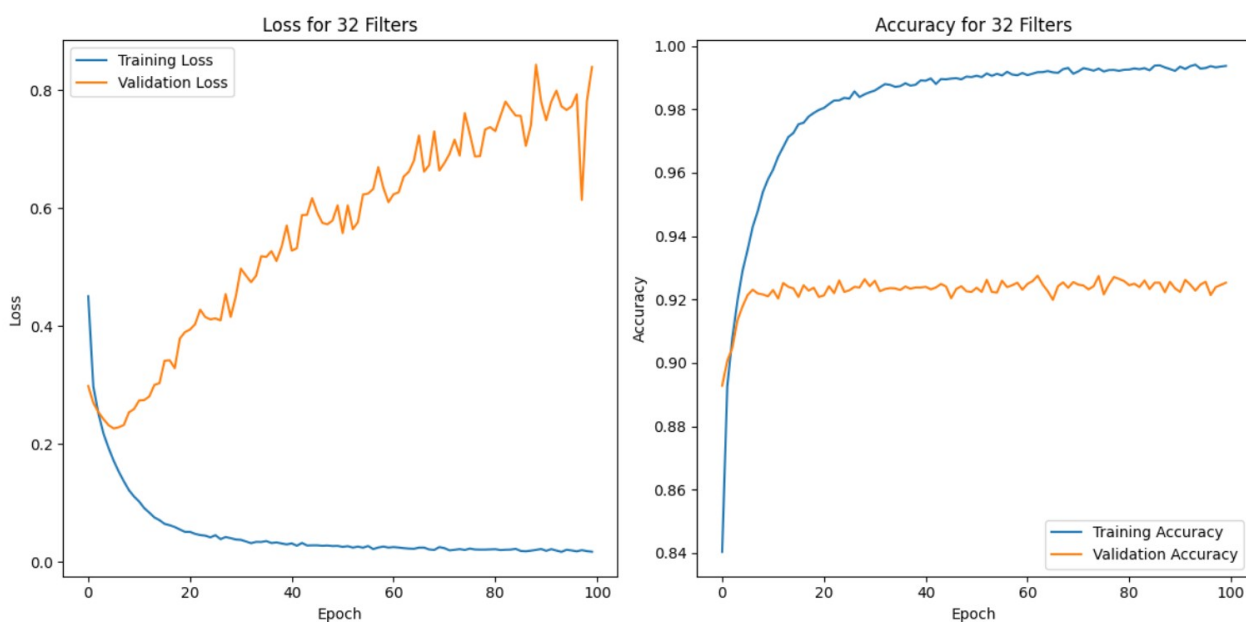
Αύξηση της αποκεντρωτικότητας: Ένας αυξημένος αριθμός φίλτρων συνεπάγεται περισσότερες ανεξάρτητες χαρακτηριστικές αναπαραστάσεις για κάθε συνελικτικό επίπεδο.

Αύξηση της Υπολογιστικής Πολυπλοκότητας: Αν και η αύξηση του αριθμού φίλτρων μπορεί να βελτιώσει την ακρίβεια του μοντέλου, μπορεί επίσης να αυξήσει την υπολογιστική πολυπλοκότητα. Αυτό οδηγεί σε μεγαλύτερο χρόνο εκπαίδευσης και μεγαλύτερες απαιτήσεις υπολογιστικού χώρου.

Κατά συνέπεια, ο αριθμός των φίλτρων είναι ένας σημαντικός παράγοντας που πρέπει να λαμβάνεται υπόψη κατά την κατασκευή του CNN. Παρακάτω φαίνονται τα διαγράμματα της συνάρτησης απώλειας και ακρίβειας και για τα δύο σύνολα για 16 και για 32 φίλτρα (εκπαίδευσης και ελέγχου).



Σχήμα 7



Σχήμα 8

Παρατηρούμε ότι για μεγαλύτερο αριθμό φίλτρων “τρέχει” περισσότερη ώρα αλλά η τιμή της ακρίβειας και η τιμή της συνάρτησης απώλειας έχουν μικρή απόκλιση για 16 και 32 πλήθος φίλτρων.