

# 1. Hyper-parameter Tuning

## Initial parameters

activation = 'tanh' | learning\_rate = 0.001 | epochs = 10 | layers = 256 – 256

accuracy: 0.9371 - loss: 0.2228 - val\_accuracy: 0.9335 - val\_loss: 0.2258

## Increasing the number of layers

activation = 'tanh' | learning\_rate = 0.001 | epochs = 10 | layers = 256 – 256 - 256

accuracy: 0.9393 - loss: 0.2119 - val\_accuracy: 0.9351 - val\_loss: 0.2189

## Increasing the number of Neurons per layer

activation = 'tanh' | learning\_rate = 0.001 | epochs = 10 | layers = 1024 – 512

accuracy: 0.9656 - loss: 0.1322 - val\_accuracy: 0.9493 - val\_loss: 0.1703

## Reducing epochs

activation = 'tanh' | learning\_rate = 0.001 | epochs = 5 | layers = 1024 – 512

accuracy: 0.9415 - loss: 0.2090 - val\_accuracy: 0.9378 - val\_loss: 0.2148

## Reducing learning rate

activation = 'tanh' | learning\_rate = 0.0005 | epochs = 10 | layers = 256 – 256

accuracy: 0.9240 - loss: 0.2764 - val\_accuracy: 0.9179 - val\_loss: 0.2815

## Increasing learning rate

activation = 'tanh' | learning\_rate = 0.01 | epochs = 10 | layers = 256 – 256

accuracy: 0.9081 - loss: 0.3044 - val\_accuracy: 0.9101 - val\_loss: 0.2948

## Selecting a different activation function

activation = 'relu' | learning\_rate = 0.001 | epochs = 10 | layers = 256 – 256

accuracy: 0.9929 - loss: 0.0308 - val\_accuracy: 0.9593 - val\_loss: 0.2094

## 2. Improvements

I experimented with modifying the network architecture by adding a dropout layer after each dense layer. This adjustment stabilized performance but did not improve accuracy compared to the original model

activation = 'tanh' | learning\_rate = 0.001 | epochs = 10 | layers = 256 – Dropout – 256 - Dropout

accuracy: 0.8946 - loss: 0.3452 - val\_accuracy: 0.9271 - val\_loss: 0.2440

### 3. Ερωτήσεις Κατανόησης:

*a. Θεωρείτε πως τα δεδομένα της MNIST είναι καλά για την εκπαίδευση ενός μοντέλου?*

*Αιτιολογείστε.*

Ναι, τα δεδομένα της MNIST είναι κατάλληλα για την εκπαίδευση ενός μοντέλου Νευρωνικού Δικτύου. Αποτελούνται από μεγάλο πλήθος εικόνων, που είναι αρκετό για τη σωστή εκπαίδευση του μοντέλου. Συνεπώς, αποτελούν ένα καλό παράδειγμα, ευκολονόητο για αρχάριους στη Βαθιά Μάθηση.

*b. Θεωρείτε πως όλα τα pixel είναι σημαντικά για την πρόβλεψη της κλάσης ενός ψηφίου?*

Τα pixels που βρίσκονται στην περίμετρο της εικόνας δε θεωρούνται σημαντικά για την πρόβλεψη. Αυτά που είναι σημαντικά, βρίσκονται πλησιέστερα στο κέντρο της εικόνας.

*c. Σε ποιες περιπτώσεις είναι καλή ιδέα να χρησιμοποιηθούν Βαθιά Νευρωνικά Δίκτυα?*

Στις περιπτώσεις που χρειαζόμαστε κατηγοριοποίηση ή λήψη αποφάσεων σε δεδομένα τα οποία δεν είναι γραμμικά. Επιπλέον, δε μας ενδιαφέρει η λογική πίσω από τη λήψη της απόφασης παρά μόνο η ακρίβεια των αποτελεσμάτων.

*d. Η Βαθιά Μάθηση μπορεί να χρησιμοποιηθεί και στους 3 κλάδους της Μηχανικής Μάθησης? (Supervised Learning, Unsupervised Learning και Reinforcement Learning)*

Ναι, η Βαθιά Μάθηση μπορεί να προσαρμοστεί ώστε να εφαρμοστεί σε όλους τους κλάδους της Μηχανικής Μάθησης.