

## ΕΡΓΑΣΙΑ 4η στα Νευρωνικά Δίκτυα

ΑΕΜ:4728

Ονοματεπώνυμο:Νικόλαος Τζάλλας

Επιβλέποντας Καθηγητής: Ιωάννης Βλαχάβας

### Hyper-parameter Tuning

Η Relu συνάρτηση ενεργοποίησης είναι καταλληλότερη για το παραπάνω νευρωνικό δίκτυο, το οποίο κατατάσσεται στα βαθιά νευρωνικά δίκτυα ( $f(x)=\ln(1+e^x)$ ). Ο ρυθμός μάθησης παρέμεινε σταθερός, έτσι ώστε να αποφευχθεί η υπερπροσαρμογή. Επειδή το σύνολο δεδομένων MNIST καθίσταται εύκολο σετ δεδομένων και ακόμα ευκολότερο με την προσθήκη βιβλιοθηκών τα κρυμμένα επίπεδα έγιναν 3 από 2, μόνο για πειραματισμό. Η αύξηση των επιπέδων αυτών προϋποθέτει πολύπλοκο πρόβλημα. Η υπερπροσαρμογή αποφεύγεται επιπλέον με σταδιακά μειούμενες ενότητες για συμπίκνωση πληροφορίας (256, 128, 64). Οι διαπεράσεις από το νευρωνικό πρέπει να είναι όσο το δυνατόν περισσότερες για μεγαλύτερη ακρίβεια. Αυξήθηκαν κατά 5 για να μην αυξηθεί πολύ και ο χρόνος εκτέλεσης.

### Άλλοι τρόποι βελτίωσης

Ο ένας τρόπος είναι με Dropout στο αρχικό μοντέλο που είχε η άσκηση. Ωστόσο δεν είναι αποδοτικός, διότι το σετ δεδομένων είναι εύκολο. Με αυτόν τον τρόπο το νευρωνικό δεν αποστηθίζει τα δεδομένα, αλλά τα κρίνει αξιολογικά, κλείνοντας μερικούς κόμβους του νευρωνικού για εξοικονόμηση ενέργειας, όπου χρειαστεί. Σχετικά με τις εικόνες του παραπάνω μοντέλου ο αποδοτικότερος τρόπος είναι με CNN-συνεξέλιξη. Οι εικόνες φιλτράρονται και το κάθε φίλτρο έχει συγκεκριμένες διαστάσεις. Τα φίλτρα αλληλοκαλύπτονται, με αποτέλεσμα να σκανάρεται η εικόνα. Επίσης με το pooling μειώνεται η πληροφορία από τα φίλτρα και εξάγεται αποτελεσματικότερα.

### Ερωτήσεις Κατανόησης

a. Θεωρείτε πως τα δεδομένα της MNIST είναι καλά για την εκπαίδευση ενός μοντέλου; Αιτιολογήστε

**Απάντηση:** Είναι κατάλληλα για εκπαίδευση μοντέλου και εξάσκηση. Αφορούν μικρές και απλές εικόνες που αποτελούνται από καλοζυγισμένα δεδομένα. (60.000 εκπαίδευση-10.000 δοκιμή).

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

**Απάντηση:** Όχι απαραίτητα. Τα περισσότερα είναι μαύρα-κενά.

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

**Απάντηση:** Σε πολύπλοκα και σύνθετα προβλήματα.

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

**Απάντηση:** Ναι σε όλους τους κλάδους, αφού η βαθιά μάθηση είναι υποσύνολο της μηχανικής.